

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2001, 14:37:17 ; Search time 41.92 Seconds

(without alignments)
1965.609 Million cell updates/sec

Title: US-09-502-698-1

Perfect score: 6990

Sequence: 1 MNCRRYPTPTQSELDSD.....EEAVAIKRTKGTLMVLS 1373

Scoring table:

BIOSUM62

Gap 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A-Geneseq_0601.*
1: /SIDSB/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSB/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSB/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSB/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSB/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSB/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSB/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSB/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSB/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSB/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSB/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSB/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSB/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSB/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSB/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSB/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSB/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSB/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSB/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSB/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSB/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDSB/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6990	100.0	1373	20	AAV04730
2	6977	99.8	2000	20	AAV04732
3	6921	99.0	2070	20	AAV04733
4	6914.5	98.9	2037	21	AAV53753
5	5085	72.7	1005	20	AAV04731
6	4080.5	58.4	856	21	AAV01383
7	3840	54.9	763	20	AAV04741
8	2981	42.6	1239	20	AAV04734
9	2457.5	35.2	1881	20	AAV24025
10	1501	21.5	318	20	AAV74150
11	1266	18.1	251	20	AAV04738

12	1228	17.6	272	20	AAV04739	PDZ domain-contain
13	1228	17.6	319	20	AAV04740	PDZ domain-contain
14	928	13.3	206	20	AAV74151	Human prostate tum
15	814	11.6	590	20	AAV04736	PDZ domain-contain
16	583	8.3	632	21	AAV6689	Membrane-bound pro
17	583	8.3	632	22	AAV87545	Human PRO1136, HO
18	580	8.3	632	22	AAV65212	Human PRO1136 (UNQ
19	580	8.3	632	21	AAV01385	Neuron-associated
20	509	7.3	2485	20	AAV19343	Amino acid sequenc
21	508	7.3	674	20	AAV92953	Fly transductosome
22	500.5	7.2	2466	16	AAV71498	Human protein tyro
23	500.5	7.2	2466	19	AAV5999	Intracellular prot
24	500.5	7.2	2466	21	AAV90272	Human PRP11 phosph
25	418.5	6.0	1161	21	AAV03832	Activin receptor b
26	409	5.9	1112	21	AAV03833	Activin receptor b
27	399	5.7	817	19	AAV48101	Human discs large
28	399	5.7	817	20	AAV30137	Amino acid sequenc
29	399	5.7	849	19	AAV48102	Human discs large
30	398.5	5.7	1277	21	AAV03149	Rat synaptic scaff
31	385	5.5	767	21	AAV22138	Human post-synapti
32	376	5.4	344	22	AAV55834	PDZ encoded domain
33	376	5.4	344	22	AAV57623	Human post-synapti
34	376	5.4	344	22	AAV58035	Human post-synapti
35	372	5.3	882	18	AAV34662	Partial PSD-93 pro
36	359	5.1	1113	22	AAV64404	Amino acid sequenc
37	357	5.1	1112	19	AAV73061	Rat GRIP, Rattus
38	347	5.0	744	21	AAV74410	PDZ domain-compris
39	339.5	4.9	396	18	AAV34666	Partial PSD-95 pro
40	339.5	4.9	1049	19	AAV73062	Rat GRIP2, Rattus
41	334.5	4.8	610	17	AAV94649	Human Fas-associat
42	315.5	4.5	414	22	AAV36687	Mammalian two-hybr
43	303.5	4.3	227	21	AAV53276	Human colon cancer
44	303.5	4.3	419	22	AAV64997	Human secreted pro
45	298	4.3	207	22	AAV57625	Human homolog of D

ALIGNMENTS

RESULT 1
ID AAV04730 standard; Protein; 1373 AA.
XX AC AAV04730;
XX DT 06-JUL-1999 (first entry)
XX DE Protein containing PDZ domain from clone 38-2-1.
XX DE (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX DE Funahashi S, Miyata S;
XX DE WPI: 1999-167423/14;
XX DE N-PSDB: AAV29908.
XX PT Protein containing PDZ domain, whose expression is enhanced by TNF stimulation - plays an important role in protein/protein

PT Interactions and is used for screening for proteins for use in
 treatment of cell proliferation disorders such as cancer

XX PS Claim 1, Page 77-84; 240pp; Japanese.

CC This sequence represents a new protein containing a PDZ domain encoded
 CC by the clone 38-2-1, whose expression in human umbilical vascular
 CC endothelial cells (HUVEC) is enhanced by stimulation with tumour necrosis
 CC factor (TNF) alpha. The new protein is used to identify proteins which
 CC bind to it (particularly to the PDZ domain) and the genes encoding them,
 CC for use in the treatment of cell proliferation disorders such as cancer.

XX Sequence 1373 AA;

Query Match Best Local Similarity 100.0%; Score 6990; DB 20; Length 1373;
 Matches 1373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYCCRRVPTTQSELDLDLDIELETERPHVDLGEFTGSSEDEDPVLAAMDAGOSTEEV 60
 Db 1 MYCCRRVPTTQSELDLDLDIELETERPHVDLGEFTGSSEDEDPVLAAMDAGOSTEEV 60
 QY 61 QAPLAMEAGQHTLELESGKGLGFTSLIDYODPDPASTVITISLVPGIAEKDGLLP 120
 Db 61 QAPLAMEAGQHTLELESGKGLGFTSLIDYODPDPASTVITISLVPGIAEKDGLLP 120
 QY 121 GDRLMFVNDVLENSLEAEALKAGPSGTVRIGVAKPLPLSPGEGYVSKEDSEFLXP 180
 Db 121 GDRLMFVNDVLENSLEAEALKAGPSGTVRIGVAKPLPLSPGEGYVSKEDSEFLXP 180
 QY 121 GDRLMFVNDVLENSLEAEALKAGPSGTVRIGVAKPLPLSPGEGYVSKEDSEFLXP 180
 Db 121 GDRLMFVNDVLENSLEAEALKAGPSGTVRIGVAKPLPLSPGEGYVSKEDSEFLXP 180
 QY 181 HSCBAGLADKPLFRADIALVGTNDADLVDESTREPSYSPENDSIYSTQASILHGSSC 240
 Db 181 HSCBAGLADKPLFRADIALVGTNDADLVDESTREPSYSPENDSIYSTQASILHGSSC 240
 QY 241 GDGLNTGSSLPSPKVDIENSCPVLDLHMSLELYTONLEODEMTPVDISMGAS 300
 Db 241 GDGLNTGSSLPSPKVDIENSCPVLDLHMSLELYTONLEODEMTPVDISMGAS 300
 QY 301 GFTINDYTPANAIDQOXCENITVWTESHPSEVISSAELPSVLPDSAGKSEHLEQSS 360
 Db 301 GFTINDYTPANAIDQOXCENITVWTESHPSEVISSAELPSVLPDSAGKSEHLEQSS 360
 QY 301 GFTINDYTPANAIDQOXCENITVWTESHPSEVISSAELPSVLPDSAGKSEHLEQSS 360
 Db 301 GFTINDYTPANAIDQOXCENITVWTESHPSEVISSAELPSVLPDSAGKSEHLEQSS 360
 QY 361 LACNAECVWLONVSKESERTINTAKNSLGKTVYSAKDGKLVSTIHGAISRDR 420
 Db 361 LACNAECVWLONVSKESERTINTAKNSLGKTVYSAKDGKLVSTIHGAISRDR 420
 QY 421 IAGDCIISTEESTISTVNAQARALRRHSLIGPDITVTPAHEEFTISIGQOSGR 480
 Db 421 IAGDCIISTEESTISTVNAQARALRRHSLIGPDITVTPAHEEFTISIGQOSGR 480
 QY 481 VMAIDFSSYTGRIPELPERESEGSESELONTAYSNMNPRLVELLRPSKSGISIV 540
 Db 481 VMAIDFSSYTGRIPELPERESEGSESELONTAYSNMNPRLVELLRPSKSGISIV 540
 QY 541 GGRMGSRSLNGEYWRGFTIKHVLDESDPAGKNGTLKPGDRIVEAPSOSESEPEKAPLCV 600
 Db 541 GGRMGSRSLNGEYWRGFTIKHVLDESDPAGKNGTLKPGDRIVEAPSOSESEPEKAPLCV 600
 QY 601 PPPPSAFAEMGSDHTOSSAKISODVDKEDDEFGYSKNTIRERGTUTGELHMLELGH 660
 Db 601 PPPPSAFAEMGSDHTOSSAKISODVDKEDDEFGYSKNTIRERGTUTGELHMLELGH 660
 QY 661 SGLGLSLAGNKRDRMSYFIVGIDPNGAAGKGRLOJADELLELNGOITLYGRSHQNASI 720
 Db 661 SGLGLSLAGNKRDRMSYFIVGIDPNGAAGKGRLOJADELLELNGOITLYGRSHQNASI 720
 QY 721 IKCAPSVYKIIIFIRNKDAVQMAVCPGNAVEPLPSNSENLONKETEPVTTSDAAYDLS 780
 Db 721 IKCAPSVYKIIIFIRNKDAVQMAVCPGNAVEPLPSNSENLONKETEPVTTSDAAYDLS 780
 QY 781 FKNNQHLEPRDGGGLIAISEEDTISGVITIKSLTEHGVAATDGRKLVGDOIILAVDEIV 840
 Db 781 FKNNQHLEPRDGGGLIAISEEDTISGVITIKSLTEHGVAATDGRKLVGDOIILAVDEIV 840

Db 781 Fknnqhleprdggliaiseedtsgvilitksltehgvaatdgrlkvgdqiavdeiv 840
 QY 841 VGYPIEKTSILKTAKTAVKTIITAEENPDSQAVPSAAGAGCKKNSQSLMVFQSGSPE 900
 Db 841 VGYPIEKTSILKTAKTAVKTIITAEENPDSQAVPSAAGAGCKKNSQSLMVFQSGSPE 900
 QY 901 PESTINRSSTPPIAFADPATCPPIPCCEETIETISKRTGLGSIYVGSPTLGAFTIH 960
 Db 901 PESTINRSSTPPIAFADPATCPPIPCCEETIETISKRTGLGSIYVGSPTLGAFTIH 960
 QY 961 EYEBGAACKDGRUWAGDQILEVNGIDIRKATHDENLWLRQTPQRVRLTYRDEAPYKE 1020
 Db 961 EYEBGAACKDGRUWAGDQILEVNGIDIRKATHDENLWLRQTPQRVRLTYRDEAPYKE 1020
 QY 1021 EYVCDTLTIELOKKPGKGLGSIYGRNDTVFVSDIYKGIADPDRGLTLOGDQILLVNG 1080
 Db 1021 EYVCDTLTIELOKKPGKGLGSIYGRNDTVFVSDIYKGIADPDRGLTLOGDQILLVNG 1080
 QY 1081 EDVRNASQEAVALKCSLGTVTEVGRIRKAGPFSERRRPSQTSQVSEGLSFTFPLSG 1140
 Db 1081 EDVRNASQEAVALKCSLGTVTEVGRIRKAGPFSERRRPSQTSQVSEGLSFTFPLSG 1140
 QY 1141 SSTSSELESSKKNALASEIOGLRTVEKKKPTDSLGISLNGYVSPUGDVPRTIAMMHP 1200
 Db 1141 SSTSSELESSKKNALASEIOGLRTVEKKKPTDSLGISLNGYVSPUGDVPRTIAMMHP 1200
 QY 1201 TGVAAGTQKLVGRIVYTCSTEGMTHTQAVMLLNKASGISIEQVYAGGVGVTGHH 1260
 Db 1201 TGVAAGTQKLVGRIVYTCSTEGMTHTQAVMLLNKASGISIEQVYAGGVGVTGHH 1260
 QY 1261 QEPASSLSFTGLTSTIFDODLGPPOCKSTLTERGPDGLGFSIVGYSRPHGLPIYVK 1320
 Db 1261 QEPASSLSFTGLTSTIFDODLGPPOCKSTLTERGPDGLGFSIVGYSRPHGLPIYVK 1320
 QY 1321 TVPKKGAASEGRKLRGDOILAVNGOSLEGVTHEEAVALLKRTGCTIMWLIS 1373
 Db 1321 TVPKKGAASEGRKLRGDOILAVNGOSLEGVTHEEAVALLKRTGCTIMWLIS 1373

RESULT 2
 ID AAY04732 standard; Protein: 2000 AA.
 XX AAY04732;
 AC 06-JUL-1999 (first entry)
 DT 06-JUL-1999 (first entry)
 DE Protein containing PDZ domain from clone 38-2-1a.
 XX PDZ domain; gene expression; human umbilical vascular endothelial cell;
 KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding;
 KW cell; proliferation disorder; cancer.
 XX Homo sapiens.
 OS Homo sapiens.
 PN WC0907846-A1.
 PD 18-FEB-1999.
 XX 12-AUG-1998; 98WO-JP03603.
 PF 19-JUN-1998; 98JP-0189944.
 PR 12-AUG-1997; 97JP-0230356.
 XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
 PA Funahashi S, Miyata S;
 PI WPI: 1999-167423/14.
 DR N-PSDB; AAX29908.
 XX Protein containing PDZ domain, whose expression is enhanced by TNF
 PT stimulation - plays an important role in protein/protein

Interactions and is used for screening for proteins for use in treatment of cell proliferation disorders such as cancer

Claim 1: Page 158-167; 240pp; Japanese.

This sequence represents a new protein containing a PDZ domain encoded by the clone 38-2-1b, whose expression in human umbilical vascular endothelial cells (HUVEC) is enhanced by stimulation with tumour necrosis factor (TNF) alpha. The new protein is used to identify proteins which bind to it (particularly to the PDZ domains) and the genes encoding them, for use in the treatment of cell proliferation disorders such as cancer.

Sequence 2000 AA:

Query Match 99.8%; Score 6977; DB 20; Length 2000;
Best Local Similarity: 99.9%; Pred. No. 0;
Matches 1371; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 MYCCRRYPTPTQSELDSDLCDELTEKPHVDLGEFGSSSTEDPVLAMTAGSTERY 60
DB 628 mvcrrtvpptqseidsldcldektekphvdigeifgsepedpvlamttagsteev 687
QY 61 QAPLAMEAGIOHIELEKSGKLGFSILDYODIPASTVITIRSLVPGIAEKDRLLP 120
DB 668 qaplmameagiohinelekygsklgfsildyqdpipastvilitrlvpgiaekdgrllp 747
QY 121 GDRLMFVNDVNLNSLEBAVEALGAPSGTVIGVAKRPLSPGEGYSAKEDSFLYPP 180
DB 748 gdrlmfvndvnlensleebavealgapsgtvirgavakrplspgegyakedsflypp 807
QY 181 HSCERAGLADKRLFRADLALVGTNDADLVDESTFSPYSPENDSTYRTASTLSLHGSSC 240
DB 808 hsceragladkrlfradlaltvgtndadvdestfspysspendsystaslslhgsac 867
QY 241 GGLINVGSSLPSPPKVDIENSCDPIVLDHMSLEELYTONLERDENTPVDISMGKPS 300
DB 866 gglinvgsslpsspkkvdienscdpivldhmsleelytonlerdentpvdismgkps 927
QY 301 GTTINDYTPANAIEOYECEENTIVTESHLPESEVTSASLPVLPDSAKGSEHLLEQSS 360
DB 928 gttindytpanaieoyeceentivteshlpesevtsaslpvlpdsakgsehllqess 987
QY 361 LACNACVWLQNVSKESFERTINIAKGNSSLGMTVSANKDGMATYRSIHGCAISRDR 420
DB 988 lacnacvwlqnvskesfertiniakgnsslgmtvsankdgmtyrsihgcaisrdr 1047
QY 421 IAGDCIISINEESTISVTNAQARALRRHSLIGPDKITTYPAEHLLEFKISLGOQSGR 480
DB 1048 iagdciiisineestisvtnaqaralrrhsligpdkittypaehllefkislgqsggr 1107
QY 481 VVALDIFSSYTGRIPELPEREGEGESELONTATSMNNORRRLVLMKPEKSLGISTY 540
DB 1108 vvaldifssytgripelperegegeeselnataysnnqrrylwrepksxslgisl 1167
QY 541 GGRKGMSRLNCGEVMGIRIKHLEDPSPAGKNGTKLPGRIVYAPQSESEPEKAPLCV 600
DB 1168 ggrkgmsrlncgevmgirikhle d pspagkngtklpgri v y a p q s e s e p e k a p l c v 1227
QY 601 PPPPSAFAMGSDHTQSSASKISQDVDKEDDEFGYSWKNIIRERYGLTGLHMLEKCH 660
DB 1228 ppppsafamgshdhtqssaskisqdvdkeddefgyswknirerygltlghlmleekh 1287
QY 661 SGLGSLACNKRBSRMKSVITVIGIDPNGAAGKGRLOIADLELTNQIILXGRSHQASST 720
DB 1288 sglgslacnkrbsrmksvityigidpngaagkgrloiadellelnqiilygrshqasst 1347
QY 721 ITCAPSKVITIFIRNKDAVNOMAVCPGNVPEPISNSENLONKETPTVTTSPAADLSS 780
DB 1348 itcapskvityifirnkda v n o m a v c p g n v p e p i s n s e n l o n k e t p t v t t s p a a d l s s 1407
QY 781 FKNVQHLLEPKDGGIGIAISEEDTLGVIKSLTEHGVAATDGRUKVGDQILAVDEIV 840
DB FKNVQHLLEPKDGGIGIAISEEDTLGVIKSLTEHGVAATDGRUKVGDQILAVDEIV 840

```

```

DB 1408 fknvqhllpkdggigiaiseedtlsgviksltehgvaatdgrlukvgdqilavdeiv 1467
QY 841 VGYPIEKRTISLKTAKMTVKTLTHAENPDSQAVPSAAGASEKKNSSQSLMVPQSGSPE 900
DB 1468 vgypiekrtisltkakmtvktlthae n p d s q a v p s a a g a s e k k n s s q s l m v p q s g s p e 1527
QY 901 PESINTSRSSPFAFADPAPCPPIPCETTIEISKRTGIGLSIVGSGDPLTGAFIIF 960
DB 1528 pesintsrsspfa f a d p a p c p p i p c e t t i e i s k r t g i g l s i v g s g d p l t g a f i i f 1587
QY 961 EYEEGACAKDGRWLWAGDOILEVNGIDRLKATHDEAINVLRQTPQVRVLTLYRDEAPYKE 1020
DB 1588 eyee g a c a k d g r w l w a g d o i l e v n g i d r l k a t h d e a i n v l r q t p q v r v l t l y r d e a p y k e 1647
QY 1021 EEVCDTLTIELOKRRKGKGLSIVCKRNDTGVFVSDIYKGGIADPDGRILIGDQILVNG 1080
DB 1648 ee v c d t l t i e l o k r r k g k g l s i v c k r n d t g v f v s d i y k g g i a d p d g r i l i g d q i l v n g 1707
QY 1081 EDVRNASQEAVALIKCSIGTWTLEVGRKAGPFHSERPQSOTQVSEGLSSTFFPLSG 1140
DB 1708 edvrnasqea v a l i k c s i g t w t l e v g r k a g p f h s e r p q s o t q v s e g l s s t f f p l s g 1767
QY 1141 SSTSESSSSKKNALASEIOLGRTVEKKKGPDSLGISIAGVSPGLDVPFIAMMHP 1200
DB 1768 s t s e s s s s k k n a l a s e i o l g r t v e k k g p d s l g i s i a g v s p g l d v p f i a m m h p 1827
QY 1201 TGVAATOKRLRGDRIVTICGSTGEMHTQAVNLKKNASGSIEMQVYAGSVVYTGHH 1260
DB 1828 t g v a a t o k r l r g d r i v t i c g s t g e m h t q a v n l k k n a s g s i e m q v y a g s v v y t g h h 1887
QY 1261 QEPASSISFTGTLTSTISIPDOLGPPQCKSTLERKPDGLGFSIYGVGSPHGDLPYVK 1320
DB 1888 q e p a s s i s f t g t l t s t i s i p d o l g p p q c k s t l e r k p d g l g f s i y g v g s p h g d l p y v k 1947
QY 1321 TVFAKGAASEDRKLRGDOIIVANGOSLEGVTHEBAVALIKFKTKGTVTLMVLS 1373
DB 1948 t v f a k g a a s e d r k l r g d o i i v a n g o s l e g v t h e b a v a l i k f k t k g t v t l m v l s 2000

```

RESULT 3

AA104733
ID AA104733 standard; Protein; 2070 AA.

XX AA104733;

AC AA104733;

DE 06-JUL-1999 (first entry)

XX Protein containing PDZ domain from clone 38-2-1b.

XX PDZ domain; gene expression; human umbilical vascular endothelial cell;

XX HUVEC; stimulation; tumour necrosis factor; TNF; protein binding;

XX cell; proliferation disorder; cancer.

XX Homo sapiens.

XX WO9907846-A1.

XX PD 18-FEB-1999.

XX PF 12-AUG-1998; 98WO-TP03603.

XX PR 19-JUN-1998; 98JP-0189944.

XX PA 12-AUG-1997; 97JP-0230356.

XX PI (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.

XX Funahashi S, Miyata S;

XX MPI: 1999-167423/14.
XX N-PSDB; AA129910.
XX Protein containing PDZ domain, whose expression is enhanced by TNF stimulation - plays an important role in protein/protein

PT Interactions and is used for screening for proteins for use in
 PT treatment of cell proliferation disorders such as cancer

PS Claim 1; Page 167-176; 240pp; Japanese.

CC This sequence represents a new protein containing a PDZ domain encoded
 CC by the clone 38-2-1b, whose expression in human umbilical vascular
 CC endothelial cells (HUVEC) is enhanced by stimulation with tumour necrosis
 CC factor (TNF) alpha. The new protein is used to identify proteins which
 CC bind to it (particularly to the PDZ domains) and the genes encoding them,
 CC for use in the treatment of cell proliferation disorders such as cancer.

XX Sequence 2070 AA:

Query Match 99.08; Score 6921; DB 20; Length 2070;
 Best Local Similarity 94.98; Pred. No. 0;
 Matches 1369; Conservative 1; Mismatches 3; Indels 70; Gaps 1;

QY 1 MWCRRTPPTTOSLSDLDLCLDELTEKPHVDGEFISGSETPVLAAMDAGQSTEEV 60
 DB mwccrrtppttgseldsldclcltelkphvdlgefsgsepedvplamtdagqsteev 687
 QY 61 QAPLAMEAGIQHTELEKSGGLGFLSDYQDPIDPASTVITIRSLVPGGIAEKDGRLLP 120
 DB qaplameagiqhileksggllgflsldyqdpidpastvilitrslvpggiaekdgrllp 747
 QY 688 qaplammeagiqhileksgkylgflsldyqdpidpastvilitrslvpggiaekdgrllp 180
 DB 688 qaplammeagiqhileksgkylgflsldyqdpidpastvilitrslvpggiaekdgrllp 747
 QY 121 GDRIMFNDVNLNLSLEAEALKGAPESTGVRIGVAKPLPLSPGEGVSAKEDSFYYP 180
 DB 121 gdrimfndvnlensleaealkgapsqtrivgvakplplspegvsaekedsflypp 807
 QY 748 gdrimfndvnlensleaealkgapsqtrivgvakplplspegvsaekedsflypp 807
 DB 748 gdrimfndvnlensleaealkgapsqtrivgvakplplspegvsaekedsflypp 807
 QY 181 HSCBEGALADKPLFRADLALVGTNDADLVDSFTESPYSPENDSIYSTQASISLHGSSC 240
 DB hscbeagaladkplfradlaltvgtndadlvdsftespyspendsiytqasislhgscc 867
 QY 808 hscbeagaladkplfradlaltvgtndadlvdsftespyspendsiytqasislhgscc 867
 DB 808 hscbeagaladkplfradlaltvgtndadlvdsftespyspendsiytqasislhgscc 867
 QY 241 GDGYNVSSSPSPKPVIVENSQCPVLDMHSLBELTQNLNLEQDNTSVDSMKRPAS 300
 DB 241 gdgynvssspspkpvivensqcpvlldmhsleltonlneqodntsvdismkrapas 927
 QY 868 gdgynvssspspkpvivensqcpvlldmhsleltonlneqodntsvdismkrapas 927
 DB 868 gdgynvssspspkpvivensqcpvlldmhsleltonlneqodntsvdismkrapas 927
 QY 301 GFTINDTPNNAIHOQECENTIWTESHLPSEYISSAELPSVLPDSAGKSEHLBOSS 360
 DB 301 gftindtpnnaihoeceentiwteshlpseyissaelpsvlpdsagksehlbooss 987
 QY 928 gftindtpnnaieqceentlwteshlpseyissaelpsvlpdsagksehlbooss 987
 DB 928 gftindtpnnaieqceentlwteshlpseyissaelpsvlpdsagksehlbooss 987
 QY 361 LACNAECVMLQNVSKESFERTINIAKGNSSLGMTVSANKDGLMIVRSIIHGAIISBDR 420
 DB 361 lacnaecvmlqnvskesfertiniakgnsslgmtvsankdglmivrsiihgaisbdr 1047
 QY 988 lacnaecvmlqnvskesfertinlakgnslgmtvsankdglmivrsiihgaisbdr 1047
 DB 988 lacnaecvmlqnvskesfertinlakgnslgmtvsankdglmivrsiihgaisbdr 1047
 QY 421 IATGDCILSINSESTISVTNQAARARHSLIGPDIKITVPAHEHEEKISLGGQSGR 480
 DB 421 iatgdcilsinsestisvtnqaararhsligpdikityvpaheheekislgqsgsr 1107
 QY 1048 iatgdcilsinsestisvtnqaararhsligpdikityvpaheheekislgqsgsr 1107
 DB 1048 iatgdcilsinsestisvtnqaararhsligpdikityvpaheheekislgqsgsr 1107
 QY 481 VMAIDIFSSTGRDIPELPEREEGEGESELQNTAYSNMNPORRELMPREPSKLSGISIV 540
 DB 481 vmaidifsstgrdipelpereegegeeselqntaysnmnporelmprepsklsgisiv 1167
 QY 1108 vmaidifsstgrdipelpereegegeeselqntaysnmnporelmprepsklsgisiv 1167
 DB 1108 vmaidifsstgrdipelpereegegeeselqntaysnmnporelmprepsklsgisiv 1167
 QY 541 GGRMGSRSLNNGEYMGKIFIKHVLNEDSPACKNGTLKPDRIIVF----- 583
 DB 541 ggrmgsrslngeymgkifikhvlnedspackngtlkpdriivf----- 1227
 QY 1168 ggrmgsrslngeymgkifikhvlnedspackngtlkpdriivf----- 1227
 DB 1168 ggrmgsrslngeymgkifikhvlnedspackngtlkpdriivf----- 1227
 QY 584 -----APSGSES 590
 DB 584 -----APSGSES 1287
 QY 1228 ltrkagpufvfmvgllmrprkspjllhnllypkynfstnfpadsllqldackapsqses 1287
 DB 1228 ltrkagpufvfmvgllmrprkspjllhnllypkynfstnfpadsllqldackapsqses 1287
 QY 591 EPEKAPLCSVPPPPPSFAEMGSDHNTQSSASKISQVDKEDDEFGYSWKNTIRERTGLTGE 650
 DB 591 epekaplcsvpppppsfaemgshntqssaskisqvdkeddefgyswkntirertglatge 1347
 QY 1288 epekaplcsvpppppsfaemgshntqssaskisqvdkeddefgyswkntirertglatge 1347
 DB 1288 epekaplcsvpppppsfaemgshntqssaskisqvdkeddefgyswkntirertglatge 1347
 QY 651 LHMTELEKSGSLGSLAGKNDKSRMSVFTVGIDPNSAGKDRGLQIADLELLEINGQILY 710
 DB 651 lhmteleksgslgslagkndksrmsvftvgidpnsagkdrqlqiadlelleingqily 1407
 QY 1348 lhmteleksgslgslagkndksrmsvftvgidpnsagkdrqlqiadlelleingqily 1407
 DB 1348 lhmteleksgslgslagkndksrmsvftvgidpnsagkdrqlqiadlelleingqily 1407
 QY 711 GRSNONASSIIKCAPSVKIIIFRNKDAVNQMVCPGNANAEPLPSNSENQNKETETTV 770
 DB 711 grsnonassiiicapsvkiiifrnkdavnqmvcpgnanaeplpsnsenqnketetv 770
 DB 711 grsnonassiiicapsvkiiifrnkdavnqmvcpgnanaeplpsnsenqnketetv 770

DB 1408 grshqnasllkcapsvkiiifrnkdavnqmvcpgnanaveplpsnsenqnketeptvt 1467
 QY 771 TSDAAYDLSSFKNVQHLELPRDOGLGIAISEEDTLGTVIRKSTREHGAATGRLKVG 830
 DB 771 tsdaaydlssfkvnqhlelprdoglglaisseedtlgtvirkstrehgaatgrlkvgd 1527
 QY 1468 tsdaaydlssfkvnqhlelprdoglglaisseedtlgtvirkstrehgaatgrlkvgd 1527
 DB 1468 tsdaaydlssfkvnqhlelprdoglglaisseedtlgtvirkstrehgaatgrlkvgd 1527
 QY 831 QIIAVALDEIVVYPIEKFISLTKAKTVKLTTHAENPDSQAVPSAAGASGERKNSQS 890
 DB 831 qiiavaldeivvypiekfisltkaktvklthhaenpdsqavpsaagaasgerknsqs 1587
 QY 1528 qiiavaldeivvypiekfisltkaktvklthhaenpdsqavpsaagaasgerknsqs 1587
 DB 1528 qiiavaldeivvypiekfisltkaktvklthhaenpdsqavpsaagaasgerknsqs 1587
 QY 891 LMPVQSSSPERESTRNRSRSTPAIFASDPATCIIITGCTETIEISGRGLGISTYGS 950
 DB 891 lmpvqssspersrnrstpaifasdpatciiitgctetieisgrglgistygs 1647
 QY 1588 lmpvqssspersrnrstpaifasdpatciiitgctetieisgrglgistygs 1647
 DB 1588 lmpvqssspersrnrstpaifasdpatciiitgctetieisgrglgistygs 1647
 QY 951 DTLGAFIIEVEYEGAAKDGRLMAGDQILEVNGIDLRKATHDEANLVNROTPOVRLT 1010
 DB 951 dtlgafiiheveyegaaakdgrlmagdqilevngidlrkathdeanlvnrotpovrlt 1707
 QY 1648 dtlgafiiheveyegaaakdgrlmagdqilevngidlrkathdeanlvnrotpovrlt 1707
 DB 1648 dtlgafiiheveyegaaakdgrlmagdqilevngidlrkathdeanlvnrotpovrlt 1707
 QY 1011 LYRDEAPYKEEVCDTLTITELQKKPKGKGLSLIVKRNDRGVFVSDIVKGLADPDGRLL 1070
 DB 1011 lyrdapyekeevcdtltitelqkkpkkgkglslivkrndrgvfvsdivkgladpdgrll 1767
 QY 1708 lyrdapyekeevcdtltitelqkkpkkgkglslivkrndrgvfvsdivkgladpdgrll 1767
 DB 1708 lyrdapyekeevcdtltitelqkkpkkgkglslivkrndrgvfvsdivkgladpdgrll 1767
 QY 1071 OGDQIILVNGEDVYVNAOSEAVAAALKKSLGTVLTVENGRIKAGPHSRRSQTSQVSEGS 1130
 DB 1071 ogdqiiilvngedvynaoeavaaalkkslgtvltvengrikagphsrrsqtsqvsegs 1827
 QY 1768 ogdqiiilvngedvynaoeavaaalkkslgtvltvengrikagphsrrsqtsqvsegs 1827
 DB 1768 ogdqiiilvngedvynaoeavaaalkkslgtvltvengrikagphsrrsqtsqvsegs 1827
 QY 1131 LSSEFPPLSGSTSESLSSSKKNALASEIQGLRTVEMKKGPTDSLGISIAGVGSPLGD 1190
 DB 1131 lssefpplsgstsesslsskknalaseiqglrtvemkkgptdslgisiaagvgsplgd 1887
 QY 1828 lssefpplsgstsesslsskknalaseiqglrtvemkkgptdslgisiaagvgsplgd 1887
 DB 1828 lssefpplsgstsesslsskknalaseiqglrtvemkkgptdslgisiaagvgsplgd 1887
 QY 1191 VPFIAMNHPTGVAAGTQKRLVGDRIYVTCGSTEGMTHTQAVNLKNASGSIEMOVVAG 1250
 DB 1191 vpfiamnhtpgvaagtqkrlvgdriyvtcgstegmthtqavnlknasgsiemovvag 1947
 QY 1888 vpfiamnhtpgvaagtqkrlvgdriyvtcgstegmthtqavnlknasgsiemovvag 1947
 DB 1888 vpfiamnhtpgvaagtqkrlvgdriyvtcgstegmthtqavnlknasgsiemovvag 1947
 QY 1251 GDVSVTVGHQHOEPASSLSFTGLTSTSIPODDIAGPPOCKSITLERGPDGFGFSIVGSGS 1310
 DB 1251 gdvsvtvghhoepasslsftglststsi poDDIAGPPOCKSITLERGPDGFGFSIVGSGS 2007
 QY 1948 gdvsvtvghhoepasslsftglststsi poDDIAGPPOCKSITLERGPDGFGFSIVGSGS 2007
 DB 1948 gdvsvtvghhoepasslsftglststsi poDDIAGPPOCKSITLERGPDGFGFSIVGSGS 2007
 QY 1311 PHGDLPIYKTVFAKGAASEDGRRLKRGDQIIAANGQSLSEGVTHEEVAVALIKRTKGTTL 1370
 DB 1311 phgdlpiyktvfkagaaesdgrlkrigdqiiaangqslsegvtheevaavalikrtkgttl 2067
 QY 2008 phgdlpiyktvfkagaaesdgrlkrigdqiiaangqslsegvtheevaavalikrtkgttl 2067
 DB 2008 phgdlpiyktvfkagaaesdgrlkrigdqiiaangqslsegvtheevaavalikrtkgttl 2067
 QY 1371 VLS 1373
 DB 1371 vls 2070
 DB 2068 vls 2070

RESULT 4

AA53753 standard; Protein; 2037 AA.

AA53753;

22-FEB-2000 (first entry)

Amino acid sequence of the MMS2 protein.

Human; MMS2; MMAC1; PDZ domain; tumour suppressor; tyrosine phosphatase;

scaffolding protein; cancer.

Homo sapiens.

WO958548-A1.

18-NOV-1999.

07-MAY-1999; 99MO-US09969.

08-MAY-1998; 98US-0084740.

(MYRI-) MYRIAD GENETICS INC.

Bartel PL, Tavtigian SV;

xx MPI: 2000-053077/04.
 DR N-PSDB: AA236453.
 xx
 PT Nucleic acids and polypeptides representing human MMS2, useful for
 detecting, diagnosing a predisposition to, and treating cancer -
 xx
 PS Claim 1, Page 93-99; 112pp; English.
 CC The present sequence represents human MMS2 protein. The MMAC1 protein
 CC binds to MMS2. The MMS2 protein has 11 post-synaptic density protein,
 CC disc-large, zo-1 (PDZ) domains and one or more of these domains
 CC interacts specifically with the carboxyl terminal amino acids of MMAC1
 CC (see AY53754). Specifically, it appears that domain 7, 10 and 13
 CC interact with MMAC1. Since MMS2 contains 11 PDZ domains and interacts
 CC with MMAC1, a known tumour suppressor having a region of homology with
 CC protein tyrosine phosphatases, MMS2 acts as a scaffolding protein in a
 CC common biological pathway with MMAC1. It is believed that the
 CC interaction between MMAC1 and MMS2 is required for the tumour suppressor
 CC activity of MMAC1. The MMS2 polypeptides, polynucleotides, fragments and
 CC specific or complex specific antibodies may be used for detecting cancer
 CC or a predisposition to cancer and screening for agents that may be used
 CC to treat MMS2 and/or MMAC1 related cancer. The polypeptides and
 CC polynucleotides may also be used to treat cancer.
 xx
 SO Sequence 2037 AA:
 Query Match 98.9%; Score 6914.5; DB 21; Length 2037;
 Best Local Similarity 96.7%; Pred. No. 0;
 Matches 1363; Conservative 7; Mismatches 3; Indels 37; Gaps 1;
 QY 1 MYCCRRTPPTQSELDLDCDIELEKPHYDLGEFTGSGSTEDPVYAMTADGSTEYV 60
 DB 628 MGCCTTTPPTQSELDLDCDIELEKPHYDLGEFTGSGSTEDPVYAMTADGSTEYV 687
 QY 61 QAPLAWMEAGIOHIELEKSGKGLGFIIDYODPIIDPASTVIRIIRSLVPGIAGKGRLLP 120
 DB 688 qaplamwsegqihieleksgkglgfslldyqdpdpastvllrlslvpgiagkgrllp 747
 QY 121 GRLMFVNDVNLNENSLLEAVNALKGAPSGYTRIGVAKPLPSPEGYVSAKEDSFLRP 180
 DB 748 grlmfvndvnlennslleavalkgapsgtyrigrvakplpspegyvsakedsflrpp 807
 QY 181 HSCERAGLADKPLFRADLALVTGNADLVDESTFSPSPENDSTYSTOASTLSHGSSC 240
 DB 808 hsceraagladrplfradlaltvgnadlvdestfespendsystoastlshgssc 867
 QY 241 GDGLNYGSSLPSPPKVDIENSCDPLVLDHMSLEELYTQNLERODENTPVDSKGPAS 300
 DB 868 gdglnygsslpsspkkvdienscdplvldhmsleelytqnllrqdentpsvdismpas 927
 QY 301 GTTINDYPPANAIEQOYECEENTVTWESHLPSEVTSASLEPSVLPDSAGSGEHLLEQSS 360
 DB 928 gttindyppanaieqoyecentvtweshlpsevtssaslepsvlpdsagsgesyhlleqss 987
 QY 361 IACNACVWLQVNVSKESFERTINIAKGNSSLTMTVSANKDGKMTVRSIRHGAISROGR 420
 DB 988 iacnacvwlqvvnvskesfertiniakgnssltmtvsankdgmvtvrsirhgaistrgr 1047
 QY 421 IAGDCILSINEESTISVTNAQARALRRHSLIGPDIKITYVPAHELEFKTSLGQOSGR 480
 DB 1048 iagdcilsineestisvtnaqaralrrhsligpdikityvpahelefktslsgqogr 1107
 QY 481 WVALDIFSSYTRGDIPELPEREGEGESELONTAYSNMNOORRRVLMPEPKSLGISIV 540
 DB 1108 wvaldifssytrgdipeleperegeeselontaysnmnoorrrvmlpepkslgisiv 1167
 QY 541 GGRGMSRLSNGEVNKGIFIKHVLSDSPAGKNGTLKPGDRIVE----- 583
 DB 1168 ggrgmssrlsngenvnkgifikhvlsdspagkngtlkpgdrivevgmdlrdshegavea 1227
 QY 584 -----APSQSESSEPEKAPLCSVPPPPSAFAEMGSDHTQSSASKI 623

DB 1228 lrkagrvvfmvqslinprapsgeesepeklpsvppppsaftaengsdhtqssaski 1287
 QY 624 SODVDKEDDEFGYSWKNIERRGTLAGELHMIIELEKSGSLGSLAGNKRDRSMVFIVGI 683
 DB 1288 sqdvkdedefgyswknierrgtlagelhmieleksgslgslagnkdrsmvflvgi 1347
 QY 684 DPNGAAGKDGRIQIDDELLETNGOILYGRSHONASSITKCAPSKYKIFIRKADVNOMA 743
 DB 1348 dpngaagkdgrlqiadelletngqillygrshonassitkcapskvklfirnkadvnoma 1407
 QY 744 VCPGNAVEPLPSNENLONKETEPTVTTSDAAVDSFKMVOHLELPKDOGLGIAISEE 803
 DB 1408 vcpгнаveplpsnennlonketeptvttstdaavdsfkmvohlelpkdoggliaisee 1467
 QY 804 DTLSCVITIKSLTBGVAAATDGRKAGDOILAVDDEIVGVPIEKRTISLTKAKMTVKLTI 863
 DB 1468 dtlsgvltksltbghvaatdgrlkvgdqlavddeivgyplekrtisltakmtvklti 1527
 QY 864 HAENDSOAVPSAAGAAGEKKNSSOSLMVPOGSPPEPSTIRNTSSRPAFAADPATC 923
 DB 1528 haendsoavpsaagaagekknssslmvpogsppepstirntssrpaafadpatc 1587
 QY 924 PIIPCETTIEISKRTGIGSLIVGSDTLGAFIIEHYEEGAACKDGRMLAGDOILEV 983
 DB 1588 plipcettieiskrtgigslivgsdttlgaftiiehyeeagaackdgrmlagdoilev 1647
 QY 984 NCIDLRKATHEAIVNROTPOVRRLTLRYDEAPRYKEEVCDTLTIELOKKRGKLGSI 1043
 DB 1648 ngidlrkatheainvrotpovrrltyrdeaprykeevcdltlietlqkpkrgkglgi 1707
 QY 1044 VGRKNDGVGVFVDIYKGGIADPDGRLIOGDOILLVNGEDVRNASEDAVAALKSGIATV 1103
 DB 1708 vgrkndgvgvfvdikyggadpdrllioqdolllvngedvrnasedaavaalksgiatv 1767
 QY 1104 LEVGRIRKAGPFHSEBRPSCQVSSEGLSSFTFPLSGSSSTSESSSKKNALASEIOGL 1163
 DB 1768 levgrirkagpfhserpsqvsseglsstfplsgsstselesskknalaseiogl 1827
 QY 1164 RTVEKKKGPSTDLSGTSINGVSSPLGDVPTFAMNHPICVANAQOKRLRGDITVTCGS 1223
 DB 1828 rtvekkkgrptdlsigtsingvssplgdvptfamnhpicyvanaqokrlrgditytcgs 1887
 QY 1224 TEGMHTQAVNLLKKNASGSIEMQVAVAGDVSVTGHHOEPASSSLFTGLTSTIFODL 1283
 DB 1888 tegmhtqavnllknasgsiemqvavagdvsvtgghoepassslftglststifodl 1947
 QY 1284 GPPCKSTLERGPDGLGFSIVGYSPPHGDLPITYKTVFAKGAASEDGRLEKRGDOITAV 1343
 DB 1948 gppckstlergpdglgfsivgysspghdplpyktvfakgaasedgrlkrqdqllav 2007
 QY 1344 NGOSLEGVTHEEAVAILKRTKCTVTLMTLS 1373
 DB 2008 ngoslegvtheeavailkrtkgtvtlmtls 2037
 RESULT 5
 ID AAY04731 standard; Protein; 1005 AA.
 AC AAY04731;
 DT 06-JUL-1999 (first entry)
 DE Mature protein containing PDZ domain from clone 38-2-1.
 XX PDZ domain; gene expression; human umbilical vascular endothelial cell;
 KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding;
 XX cell; proliferation disorder; cancer.
 OS Homo sapiens.
 XX WO907846-A1.
 PN

XX 18-FEB-1999.
 PD 12-AUG-1998; 98WO-JP03603.
 PF 19-JUN-1998; 98JP-0189944.
 PR 12-AUG-1997; 97JP-0230356.
 XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
 PA Funahashi S, Miyata S;
 PI WPI; 1999-167423/14.
 DR N-PSDB; AAX29908.
 XX Protein containing PDZ domain, whose expression is enhanced by TNF
 PT stimulation - plays an important role in protein/protein
 PT interactions and is used for screening for proteins for use in
 PT treatment of cell proliferation disorders such as cancer
 XX
 PS Claim 1; Page 84-87; 240pp; Japanese.
 CC This sequence represents the mature portion of a new protein containing
 CC a PDZ domain encoded by the clone 38-2-1, whose expression in human
 CC umbilical vascular endothelial cells (HUVEC) is enhanced by stimulation
 CC with tumour necrosis factor (TNF) alpha. The new protein is used to
 CC identify proteins which bind to it (particularly to the PDZ domains) and
 CC the genes encoding them, for use in the treatment of cell proliferation
 CC disorders such as cancer.
 CC
 SQ Sequence 1005 AA;

Query Match 72.7%; Score 5085; DB 20; Length 1005;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1005; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 MLQVSKSEFERTINIAKNGSSLGMTVSANKDGLMTVRSIIHGSAISRGRIGDCIL 428
 DB 1 mlgnvksesferlniakngsslgmtvsankdglmtvrsiihgalsrdgrlsgcll 60
 QY 429 SINESTISVTNAQARMLRRHSILGPDIKITYPAEHLFEKSLIGQOSRWALDIFS 488
 DB 61 sineestisvtvnagarmllrrhsilgpdikitypaehlfekslisqgsrwalidfs 120
 QY 489 SYTGRIPELPEREGGESEELONTAYSMNQPRVLELWREPSKSLGISIVGRGMSR 548
 DB 121 sytgripelperegegeeseldntaysmnqprvelwrepksisligvgrgmstr 180
 QY 549 LSNGEVMRGITIKVNLVLEDSPPAGKNGITAKPGDRIYEAPOSSEPEKAPLCSVPPPPSAF 608
 DB 181 lsngevmrgitikhvnlvledspagkngitlkpgdriyeapsqesepkaplcsvppppsaf 240
 QY 609 AEMSDHTOSSASKISODVDKEDFEFGYSMKNIRRYGTLFGLMITEKGHSLSGLSLA 668
 DB 241 aemsdhtgssaskisodvdkedefegyswmknirrygtlfglmltekyhslslgslsla 300
 QY 669 GNMKRSRMVSFIVGIDPNGAAGKDRLOIADDELLEINGOILYGRSHONASSIICARSKV 728
 DB 301 gnmkrsrmvsfivgidpngaagkdrloiaddelleingqilysrshnaslsikcarksv 360
 QY 729 KIITIRKDAVNQMAVCPGNAVEPLPSNSENLOKETEPTVYTSDAAVDLSSFKNVQHLE 788
 DB 361 kiitirkdavnqmavcpgनावेप्लप्सनसेनलोकतेप्टवत्सदावदलस्सफक्वह्ले 420
 QY 789 LPKOGGLGTAISEEDTLGTYIKSLTEHGVAARDGLRYKGDOLLAADDERIVVCPLEKF 848
 DB 421 lpkoggglgtaiseedtlsgtyiksltehgvaardglrykvgdqlaavdderivvcpylekf 480
 QY 849 ISLTKTKMKTVKLTIIHAENPDSQAVPSAAGAASGEKKNSQSLWPVPGSGPEPSINTS 908
 DB 481 isltktkmtkvltlihaenpdsqavpsaagaasgekknssqslwvpgsgpepsintts 540

QY 909 RSTPAIFASDPATCPIIPGCEITIEISKRTGTLGSLIVGSDTLGALFIHEVEEGAA 968
 DB 541 rsttpaifasdpatcpilpgcettieiskrtgltgslivgstdlgaalfiheveegaa 600
 QY 969 CKDGRINAGDOILEVNGIDLRKATHDEAINVLRQTPORVRLTLRDEAPYKEEVCPTLT 1028
 DB 601 ckdgrlwnagdqilevngidlrlkathdeainvrlrtpqvrltlrdeapykeevcdltt 660
 QY 1029 TELKKKGKGLSLIVKRNDRGTGVFSDIYKGIADDDGRILQGDOLLVNGEDVRNAAQ 1088
 DB 661 telkkkgkglslivkrndrgtgvfsvdiykgiadddgrilqgdolllvngedvrnaaq 720
 QY 1089 EAVALLKCSLGTIVTEVGRKACPFHSERRPSQTSQVSEGLSFFFPPLSGSSTSESL 1148
 DB 721 eavaallkcslgtivtevgrikacpfhserrpsqtsqvseslsfffpplsastsesle 780
 QY 1149 SSKKNALASFTGRLVRYEAKKGPDSGLSISIAGVSPGLADVFIFLMMHPGVAAQTO 1208
 DB 781 sskknalaesftgrlvryeakkgpdsdglslisagiavspgladviflammhpgvvaqto 840
 QY 1209 KLRYGDRIVTICGSTEGMTHTQAVNLKKNASGSIEMQVAGGVSVYTGHHOBPASSL 1268
 DB 841 klrygdrivtictgstegmthtqavnlkknasgsiemqvaggvsvytgghobpassl 900
 QY 1269 SFTGLTSTSIPODDLGPPOCKSTILREKPGDLGFSIVGYGSPHGLPIYKTVFARGAA 1328
 DB 901 sftglststsiipoddlgpockstilerkpgdldgfsivgygspghlpiykvtfargaa 960
 QY 1329 SEDGRILKRGQIITAVNGOSLEGVTHEEAVAILKTKCTVTLMLVLS 1373
 DB 961 sedgrilkrqdiitavngoslegvtheeavailkktctvtlmlvls 1005

RESULT 6
 AAB01383
 ID AAB01383 standard; Protein; 856 AA.
 AC AAB01383;
 XX 20-OCT-2000 (first entry)
 DE Neuron-associated protein.
 XX
 KW Neuron associated protein; NEUAP; neurological disorder; epilepsy;
 KW ischemic cerebrovascular disease; stroke; cerebral neoplasm;
 KW Alzheimer's disease; Pick's disease; Huntington's disease;
 KW dementia; Parkinson's disease; demyelinating disease; meningitis;
 KW prion disease; kuru; Creutzfeldt-Jakob disease; neurofibromatosis;
 KW cerebral palsy; muscular dystrophy; central nervous system; CNS;
 KW peripheral nervous system; PWS; myopathy; schizophrenia;
 KW actinic keratosis; arteriosclerosis; atherosclerosis; burns;
 KW cirrhosis; hepatitis; mixed connective tissue disease; MCMD;
 KW myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;
 KW autoimmune disease; inflammation; acquired immunodeficiency syndrome;
 KW AIDS; Addison's disease; adult respiratory distress syndrome;
 KW allergy; ankylosing spondylitis; amyloidosis; anaemia; asthma;
 KW Werner syndrome; trauma; human.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 1..65
 FT Modified-site /label= PDZ domain signature
 FT Modified-site 18
 FT Modified-site /note= "Potential glycosylation site"
 FT Modified-site 20
 FT Modified-site /note= "Potential phosphorylation site"
 FT Modified-site 69
 FT Modified-site /note= "Potential phosphorylation site"
 FT Modified-site 71
 FT Modified-site /note= "Potential phosphorylation site"
 FT Modified-site 73
 FT Modified-site /note= "Potential phosphorylation site"


```

Dh 121 knlerygltlgelhmlielekghsglglslagnkdrsmvflvgldpngaagkgrlgl 180
Qy 698 ADELLEINGQILLYGRSHONASSIIKCAPSKYKIFIRNKADVNOMACPGNAVEPLDSNS 757
Dh 181 adelieingqllygrshonassiiikcapskvklifirnkadvnmacpgnavaplsns 240
Qy 758 ENLONKETEPTVTSSDAVDLSSEFNVOHLELPKDOGLGIAISEEDTLSCGVIKSTLER 817
Dh 241 enlqketepvttsdaavdlssfkvghlelphdggglglalsedtlsgvllkslteh 300
Qy 818 GVAATDGLKLVGDDIILAVDEIVVGYPIEKFTSLTKAKMTVKITIAENPDSQAVSAA 877
Dh 301 gvaatdglkvlvgddilavdeivvgyplekftsltkakmtvkltiaenpdsqavpaa 360
Qy 878 GAASGEKKNSSOSLAMPQSGSPESIRNTSRSTPAIFADPACPIIPCCETTIEISK 937
Dh 361 gaasgeekknasqslmvpqsgspesirntsrstpaifadpactpiipccettietisk 420
Qy 938 GRTGLGSIYGGSDTLGAFITHEVYEEGAACKDGRLMAGDQILEVNGIDLRKATHDEAI 997
Dh 421 grtglglsiyggsdttlgaftihevyeeagaackdgrlmagdqillevngidlrkathdeai 480
Qy 998 NVLROTORVRLTLTYRDEAPYKEEVCDDTLTIELQKKPKGGLGIYIGKRNDRGVFVSDI 1057
Dh 481 nvlrotoorvrltlyrdeapykeevecdtltelqkkpkgyglstlygkndrgvfvdsi 540
Qy 1058 VKGSIADPDGRLTIGDDIILVNGEDVRNASQEAVALKLSGLTVEVGRIRKAGPFHSE 1117
Dh 541 vkgsiadpdgrltimgddillvngedvrnatqevaallkkslgtvlelvgrikagpfhse 600
Qy 1118 RRPQOTQVSESGSLSTFTPLSGSSTSESLSSSKKNAALASEIOGLRTVEKKKPTTSLG 1177
Dh 601 rrpqotqvsegsllstftplsgsstselessskknalaseioglrtvekkkpttsgl 660
Qy 1178 ISIGVGSPGLGDPVPIFIAMMHPTGVAOQKRLVRDRIYITIGSTEGMHTHQAVALMLK 1237
Dh 661 isigvgspglgdvpiifiamhptgvaqcklrvgrlyitlgtsegmhthcvaanllk 720
Qy 1238 NASGSIEMOVVAGDVSVTGHQEPASSLSLFTGLTSTISFDODLGPPOCKSTTLERGP 1297
Dh 721 naegslsemvvaagdvsvtghqepasslsftglststisfdodlppocksttlergp 780
Qy 1298 DGLGFSITVGGSPHGLPIYKTVFPAKGAASEGRKLRDQIILAVNGOSIEGTHHEAV 1357
Dh 781 dglgfsitvgygspghdpliykvtvfkagaasegrlkrdqillavngosiegtthheav 840
Qy 1358 AILKRTKGTVMVLS 1373
Dh 841 ailkrktgtvtmlvls 856

```

RESULT 7
AAV04741
ID AAV04741 standard; Protein: 763 AA.

XX AAV04741:

XX 06-JUL-1999 (first entry)

XX PDZ domain-containing protein gene encoded by clone 32-8-1/5R3.

XX PDZ domain: gene expression; human umbilical vascular endothelial cell;

XX HOVEC; stimulation; tumour necrosis factor; TNF; protein binding; PCR;

XX cell; proliferation disorder; cancer; primer; amplification.

XX Homo sapiens.

XX MO9907846-A1.

XX 18-FEB-1999.

XX 12-AUG-1998: 98MO-JP03603.

```

PR 19-JUN-1998: 98JP-0189944.
PR 12-AUG-1997: 97JP-0230356.
XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX Funahashi S, Miyata S;
XX MPI: 1999-167423/14.
DR N-PSDB: AAX29960.
XX
PT Protein containing PDZ domain, whose expression is enhanced by TNF
PT stimulation - plays an important role in protein/protein
PT interactions and is used for screening for proteins for use in
PT treatment of cell proliferation disorders such as cancer
XX
PS Example 2; Page 127-134; 240pp; Japanese.
XX
CC This sequence represents a new protein containing PDZ domains encoded
CC by clone 32-8-1/5R3, whose expression in human umbilical vascular
CC endothelial cells (HVEC) is enhanced by stimulation with tumour necrosis
CC factor (TNF) alpha. The new protein is used to identify proteins which
CC bind to it (particularly to the PDZ domains) and the genes encoding them,
CC for use in the treatment of cell proliferation disorders such as cancer.
XX
SQ Sequence 763 AA:

```

Query Match 54.9%; Score 3840; DB 20; Length 763;
Best Local Similarity 99.7%; Pred. No. 1.5e-301;
Matches 761; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

Qy 611 MGSDDTQSSASKISQDVYKDEDFEYSWKNIERYGTLTGLHMLEKSHGSLSLAGN 670
Dh 1 mgsdhtqssaskisqdvkydeefywknierygtltglhmllekshgslslagn 60
Qy 671 KDRSRMSVFIVGIDPNKAAGKDRQLQADELEINGQILYGRSHONASSIIKCAPSKYK 730
Dh 61 kdrsrmsvflvgldpnaagkdrqlqadelleingqllygrshonassiiikcapskvkl 120
Qy 731 IFIRNKDVAQNOMAVCPNNAVEPLPSNSENLONKTEPTVTSSDAVDLSSEFNVOHLELP 790
Dh 121 ifirnkadvnqomavcpnnavelpspnsenlonkteptvtssdaavdlssfkvghlelph 180
Qy 791 KDQGGGLAISIEDTLSGVLIKSLTEHGAATDGLKRVGQIILAVDEIVVGYPIEKFTIS 850
Dh 181 kdqggglaisiedtlsgvliksltehgaatdgrlkgvqillavdeivvgyplekftis 240
Qy 851 LKTKAKMTVKITIAENPDSQAVSAAAGSEKKNSSOSLAMPQSGSPESIRNTSR 910
Dh 241 lktkakmtvkltiaenpdsqavpaaagseekknssoslamvpqsgspesirntsr 300
Qy 911 STPAIFASDPACPIIPCCETTIEISKRTGLGSIYGGSDTLGAFITHEVYEEGAACK 970
Dh 301 stpaifasdpactpiipccettietiskrtglglsiyggsdttlgaftihevyeeagaack 360
Qy 971 DGRIMAGDQILEVNGIDLRKATHDEAINVLRQTPQRYRLTLKRDAPYKEEVCDDTLTIE 1030
Dh 361 dgrlmagdqllevngidlrkathdeainvlrtpqrvrltlydeapykeevecdtlte 420
Qy 1031 LOKKPKGGLSIYGRKNDGVSVSDIVKGIADPQSRLOGSOIILVNGEDVRNASQEA 1090
Dh 421 lqkpkpglglstlygrkndgvsvsdivkgiadpqrlllgvngedvrnasqea 480
Qy 1091 VAALLKCSLTGVYLENGRIKAGPFHSRRRQSTQVSEGLSFTPLSSSTSESLSS 1150
Dh 481 vaallkcsigtvlelvgrlkaagpfhserrrpsqsvseglstftplsgsstseless 540
Qy 1151 SKKNALASEIQLRTVEKKKPTDLSGISIAGVGSPGLGDPVPIFIAMMHPTGVAOQTK 1210
Dh 541 skknalaseiqlrtvekkkptdlsigisiagvgspglgdvpiifiamhptgvaqtkl 600
Qy 1211 RVGDRIYITIGSTEGMHTHQAVALNLKNASGSIEMOVVAGDVSVTGHQEPASSLSL 1270

```

Db 601 rvgddivlvcstegmtctgavnlknasgslemvavagdvsvltghhgpasslsf 660
 Qy 1271 TGLTSTSTFODDLGPPCKSTLRENGPDGLSPVIGVGSPPHGDPIYKTYFAKGAASE 1330
 Db 661 TGLTSTSTFODDLGPPCKSTLRENGPDGLSPVIGVGSPPHGDPIYKTYFAKGAASE 720
 Qy 1331 DGRLRGDDIIVANGOSLEGVTHEEVAVALIKRTKGTVTLMVLS 1373
 Db 721 dgrlkrqgdllavngsglegvtheeavallkrtkgtvtlmvls 763

RESULT 8

AA04734
 ID AAY04734 standard; Protein: 1239 AA.

AC AAY04734;

DT 06-JUL-1999 (first entry)

DE Protein containing PDZ domain from clone 38-2-1c.

KM PDZ domain; gene expression: human umbilical vascular endothelial cell;

KM HUVEC; stimulation: tumour necrosis factor; TNF; protein binding;

KM cell; proliferation disorder; cancer.

OS Homo sapiens.

PM WO9907846-A1.

PD 18-FEB-1999.

PF 12-AUG-1998; 98WO-JP03603.

PR 19-JUN-1998; 98JP-0189944.

PR 12-AUG-1997; 97JP-0230356.

PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.

PI Funahashi S, Miyata S;

DR WPI: 1999-167423/14.

DR N-PSDB; AAX29911.

XX Claim 1; Page 177-182; 240pp; Japanese.

CC This sequence represents a new protein containing a PDZ domain encoded

CC by the clone 38-2-1c, whose expression in human umbilical vascular

CC endothelial cells (HUVEC) is enhanced by stimulation with tumour necrosis

CC factor (TNF) alpha. The new protein is used to identify proteins which

CC bind to it (particularly to the PDZ domains) and the genes encoding them,

CC for use in the treatment of cell proliferation disorders such as cancer.

XX Sequence 1239 AA;

Query Match 42.6%; Score 2981; DB 20; Length 1239;

Best Local Similarity 99.7%; Pred. No. 1,1e-231;

Matches 581; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MYCCRRVPTPTQSLDSDLDIELEKPHVDLGEFTSSSETEDPVLAMTADAGSTERY 60
 Db 628 MYCCRRVPTPTQSLDSDLDIELEKPHVDLGEFTSSSETEDPVLAMTADAGSTERY 687
 Qy 61 GQPLAMMEGIOHIELEKSGKIGFSLDYPDPASPVTIIRSVPGGIAEKGRLLP 120
 Db 688 GQPLAMMEGIOHIELEKSGKIGFSLDYPDPASPVTIIRSVPGGIAEKGRLLP 747
 Qy 121 GDRLMFVNDVNLNLSLEAEALKGAPSGTVRGVAKPLPLSPGEGYSAKEDSFLYPP 180

Db 748 gdrlmfndvnlensleaeavalkgapsgtvrlgvakplrlspggyvakedsflypp 807
 Qy 181 HSCERAGLADKPLFRADALAVGTNDADLVDESTFSPSPSPENDSTYSTQASLSLHGSSC 240
 Db 808 HSCERAGLADKPLFRADALAVGTNDADLVDESTFSPSPSPENDSTYSTQASLSLHGSSC 867
 Qy 241 GDLNKGSSLPSPPKDVIENSCDPLDLMHSELTYNQLERODEMTPSVDISMGAS 300
 Db 868 gdlngsslpsspdkdvienscdpldlmhseeltynqlerodentpsvdismpas 927
 Qy 301 GFTINDYTPANAIEQOYECENTIVWTESHLPSEVISSEALPSLPDSAGKSGEHLLEQSS 360
 Db 928 gftindytpanaieqycentivwteshlpsevissealpslpdsagksgehllegss 987
 Qy 361 IACNAECVWLQNVSKESFERTINAKNSLQMTVSAKNDGLMTVRSIIHGASISDGR 420
 Db 988 IACNAECVWLQNVSKESFERTINAKNSLQMTVSAKNDGLMTVRSIIHGASISDGR 1047
 Qy 421 IAIIGCIIISINEESTISVTNQAARLRHSLIGDITITYPAHLEFPKISLGQOSGR 480
 Db 1048 IAIIGCIIISINEESTISVTNQAARLRHSLIGDITITYPAHLEFPKISLGQOSGR 1107
 Qy 481 VVALDIFSSYTGRIPELPEREGEGESELQNTAVSNMNOPTRELVREPSSLSGISIV 540
 Db 1108 VVALDIFSSYTGRIPELPEREGEGESELQNTAVSNMNOPTRELVREPSSLSGISIV 1167
 Qy 541 GGRGMSRLSNGEVMKGFIFKHVLEDSPAKNGTKLPKGDRIYE 583
 Db 1168 ggrgmsrlsngvmkgfifkhvledspakngtklpkgdrtive 1210

RESULT 9

AA24025
 ID AAY24025 standard; Protein: 1881 AA.

AC AAY24025;

DT 29-SEP-1999 (first entry)

DE Amino acid sequence of the human MMS1 protein.

KM Human; MMS1 protein; MMAC1 interacting protein; tumour suppression;

KM MMAC1 pathway; immunogen; cancer; cell neoplastic growth.

OS Homo sapiens.

PM WO9936566-A1.

PD 22-JUL-1999.

PF 19-JAN-1999; 99WO-US00095.

PR 20-JAN-1998; 98US-0071861.

PA (MYRI-) MYRIAD GENETICS INC.

PI Bartel PL, Tavtigian SV;

DR WPI: 1999-458472/38.

DR N-PSDB; AAX86366.

XX MMS1, an MMAC1 (tumour suppressor) interacting protein and related

XX polynucleotides

XX Claim 14; Page 88-93; 107pp; English.

CC The present sequence represents a MMS1 protein. The protein is a MMAC1

CC interacting protein which is involved in tumour suppression activity

CC in the MMAC1 pathway. MMS1, antigenic fragments or fusion proteins of

CC these are used as immunogens for antibody production. Primers derived

CC from MMS1 genomic clones can be used for identification of MMS1 genes

CC and for synthesis by amplification of MMS1 DNA or RNA. Detecting an

XX MPI: 1999-621386/54.
 DR N-PSDB; AA252969.
 XX
 PT New human nucleic acid sequences from pancreatic tumors, and related
 PT proteins -
 XX
 PS Claim 23; Page 449-450; 502pp; German.
 XX
 CC This invention describes novel polypeptides and their encoding nucleic
 CC acids derived from human pancreatic tumor tissue which have cytostatic
 CC activity. The sequences are also useful in producing pharmaceutical
 CC compositions for treatment of pancreatic tumors. AA73814-74252
 CC represent protein fragments encoded by the human pancreatic tumor CDNA
 CC library derived expressed sequence tag (EST) sequences represented in
 CC AA52858-253014.
 CC
 XX
 SQ Sequence .318 AA;

Query Match 21.5%; Score 1501; DB 20; Length 318;
 Best Local Similarity 96.7%; Pred. No. 6e-113;
 Matches 297; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 790 PKDGGGLGIAISEEDTLSCVITKSLTEHGVATDGRKLVGDQILAVDEIVVGYPIEKPI 849
 DB 9 prrggglglaisedtlsgvltksltelhgvaaatdgrlkvgdqlavddeivvgypiekti 68
 OY 850 SLTKAKMTVKLTIAHENDSQAVPSAAGASGKKNSSQSLMVPQSGPEPESTRNTSR 909
 DB 69 sltkakmtvklthnaendsqavpsaagaasgekknssqslmvpqsgpepesitrntsr 128
 OY 910 SSTPAIFASDPATCPIPCCECTTIFISKRTGLSIVGSGSTLLGAFIIVHEVGEAAG 969
 DB 129 sstpaifasdpatcpilpccecttelskyrtglslvgsstllgailihevgeaag 188
 OY 970 KDGRLMAGDQILEVNGIDLRKATHDEAINVLRQTPQRYELTYRDEAPYKEEVCDTLTI 1029
 DB 189 kdgrlmagdqilevngidlrkatheainvltqtpqrvlyltyrdeapyeecvcdtlti 248
 OY 1030 ELQKRPKGLSIVGKNDTGVFVSDIVKGIADPGRLLIGDQILLVNGEDVRNASE 1089
 DB 249 elqkrpkglslvkgndtgvfvsdlvkgiadadgrllmgdqlimvngedvrnatge 308
 OY 1090 AVAALLK 1096
 DB 309 avavwvk 315

RESULT 11
 AA04738
 ID AA04738 standard; Protein; 251 AA.

XX AC AA04738;
 XX
 DT 06-JUL-1999 (first entry)
 XX
 DE PDZ domain-containing protein gene encoded by clone FH750.
 XX
 KW PDZ domain; gene expression; human umbilical vascular endothelial cell;
 KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding; PCR;
 KW cell; proliferation disorder; cancer; primer; amplification.
 XX
 OS Homo sapiens.
 XX
 PN WO9907846-A1.
 XX
 PD 18-FEB-1999.
 XX
 PF 12-AUG-1998; 98WO-JP03603.
 XX
 PF 19-JUN-1998; 98JP-0189944.
 PR 12-AUG-1997; 97JP-0230356.

XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
 PA Funahashi S, Miyata S;
 XX
 PI
 XX
 DR MPI: 1999-167423/14.
 DR N-PSDB; AA29969.
 XX

PT Protein containing PDZ domain, whose expression is enhanced by TNF
 PT stimulation - plays an important role in protein/protein
 PT interactions and is used for screening for proteins for use in
 PT treatment of cell proliferation disorders such as cancer
 XX

PS Example 7; Page 149-151; 240pp; Japanese.

CC This sequence represents a new protein containing PDZ domains encoded
 CC by clone FH750, whose expression in human umbilical vascular endothelial
 CC cells (HUVEC) is enhanced by stimulation with tumour necrosis factor
 CC (TNF) alpha. The new protein is used to identify proteins which bind
 CC to it (particularly to the PDZ domains) and the genes encoding them,
 CC for use in the treatment of cell proliferation disorders such as cancer.
 CC
 XX
 SQ Sequence 251 AA;

Query Match 18.1%; Score 1266; DB 20; Length 251;
 Best Local Similarity 99.6%; Pred. No. 4.1e-94;
 Matches 250; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 341 PSYLPDSAGKSGEHLLEOSSLACNAECVMTQNVKSEFERTINAKGSSLGMTVSANKD 400
 DB 1 psvlpdsagksgeyllsgslacnaecvmtqnvksesfertlnakgsslgmtvsankd 60
 OY 401 GLGMITVRSIHGCAISRGRALIGDCIISINEESTISVTNMAQAMLRHSLIGPDITIT 460
 DB 61 glgmivrsilhgcaisrgraligdcilisineestisvtnmaqamlrhsligpditit 120
 OY 461 YVPAEHLPEFKTSLGOQSGRVVALDIFSSYGRDIPERPEREGGSESELONTAYSMNN 520
 DB 121 yvpaehleefkislqgsgryvaldlfssygrdipeperegeeseelqntaysmn 180
 OY 521 QPRVLEWRPESKSLGISIVGSGKSRSLNGEVARGIFIKHVLDSPPAKNGTLKPGDR 580
 DB 181 qprvlewrpkslglslvgsrgksrslngevargifikhvldspakngtlkpgdr 240
 OY 581 IVEAPSGESE 591
 DB 241 iveapsgese 251

RESULT 12
 AA04739
 ID AA04739 standard; Protein; 272 AA.

XX AC AA04739;
 XX
 DT 06-JUL-1999 (first entry)
 XX
 DE PDZ domain-containing protein gene encoded by clone FH850.
 XX
 KW PDZ domain; gene expression; human umbilical vascular endothelial cell;
 KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding; PCR;
 KW cell; proliferation disorder; cancer; primer; amplification.
 XX
 OS Homo sapiens.
 XX
 PN WO9907846-A1.
 XX
 PD 18-FEB-1999.
 XX
 PF 12-AUG-1998; 98WO-JP03603.
 XX
 PF 19-JUN-1998; 98JP-0189944.

PR 12-AUG-1997; 97JP-0230356.
XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX Funahashi S, Miyata S;
XX WPI: 1999-167423/14.
DR N-PSDB; AAX29970.
XX
XX Protein containing PDZ domain, whose expression is enhanced by TNF
PT stimulation - plays an important role in protein/protein
PT interactions and is used for screening for proteins for use in
PT treatment of cell proliferation disorders such as cancer
XX
PS Example 7; Page 152-154; 240pp; Japanese.
XX
XX This sequence represents a new protein containing PDZ domains encoded
CC by clone FH950, whose expression in human umbilical vascular endothelial
CC cells (HUVEC) is enhanced by stimulation with tumour necrosis factor
CC (TNF) alpha. The new protein is used to identify proteins which bind
CC to it (particularly to the PDZ domains) and the genes encoding them,
CC for use in the treatment of cell proliferation disorders such as cancer.
XX
SQ Sequence 272 AA;

Query Match 17.6%; Score 1228; DB 20; Length 272;
Best Local Similarity 99.6%; Pred. No. 5.6e-91;
Matches 242; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 341 PSVLPDSAGKSEHLEQSSSLACNAECVMQNSKESFERTINIAKNSISGIMTVSANKD 400
DB 1 psvlpdsagkseylllegsslacnaecvmqlqvakesfertinlaknsslgmtvsankd 60
QY 401 GLGMIVRSIIHGGAISRDGRIAGDCILISNEESTISVTNQAARMLRRHSLIGPDIKIT 460
DB 61 glgmivrsiihgaisrdgriaidcilsineestisvtnqaramlrrhsligpdkit 120
QY 461 YVPAEHLEEFKISIGQSGRYMALDIFSSYGRDIPELPEREGEGESSELQNTAYSNMN 520
DB 121 yvpaeheefkisiqsgsgrymaldifssygrdipelperegegeselqntaysnmn 180
QY 521 QPRVVELMRPEPSKSLGISIVGSGRMSRLSNGEVMRGIFIKHVLSDSPACKNGTLKPGDR 580
DB 181 qprvvelmrpepsksglgsivgsgrmsrlsngewmrgiflkhvledspakngtlkpgdr 240
QY 581 IVE 583
DB 241 lve 243

RESULT 13
AAV04740
ID AAV04740 standard; Protein: 319 AA.
XX
AC AAV04740;
XX
DT 06-JUL-1999 (first entry)
XX
DE PDZ domain-containing protein gene encoded by clone FH950.
XX
XX PDZ domain; gene expression; human umbilical vascular endothelial cell;
KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding; PCR;
KW cell; proliferation disorder; cancer; primer; amplification.
XX
OS Homo sapiens.
XX
XX WO907846-A1.
XX
XX 18-FEB-1999.
XX
XX 12-AUG-1998; 98WO-JP03603.
XX

PR 19-JUN-1998; 98JP-0189944.
XX 12-AUG-1997; 97JP-0230356.
XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX Funahashi S, Miyata S;
XX WPI: 1999-167423/14.
DR N-PSDB; AAX29971.
XX
XX Protein containing PDZ domain, whose expression is enhanced by TNF
PT stimulation - plays an important role in protein/protein
PT interactions and is used for screening for proteins for use in
PT treatment of cell proliferation disorders such as cancer
XX
PS Example 7; Page 155-158; 240pp; Japanese.
XX
XX This sequence represents a new protein containing PDZ domains encoded
CC by clone FH950, whose expression in human umbilical vascular endothelial
CC cells (HUVEC) is enhanced by stimulation with tumour necrosis factor
CC (TNF) alpha. The new protein is used to identify proteins which bind
CC to it (particularly to the PDZ domains) and the genes encoding them,
CC for use in the treatment of cell proliferation disorders such as cancer.
XX
SQ Sequence 319 AA;

Query Match 17.6%; Score 1228; DB 20; Length 319;
Best Local Similarity 99.6%; Pred. No. 7.4e-91;
Matches 242; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 341 PSVLPDSAGKSEHLEQSSSLACNAECVMQNSKESFERTINIAKNSISGIMTVSANKD 400
DB 1 psvlpdsagkseylllegsslacnaecvmqlqvakesfertinlaknsslgmtvsankd 60
QY 401 GLGMIVRSIIHGGAISRDGRIAGDCILISNEESTISVTNQAARMLRRHSLIGPDIKIT 460
DB 61 glgmivrsiihgaisrdgriaidcilsineestisvtnqaramlrrhsligpdkit 120
QY 461 YVPAEHLEEFKISIGQSGRYMALDIFSSYGRDIPELPEREGEGESSELQNTAYSNMN 520
DB 121 yvpaeheefkisiqsgsgrymaldifssygrdipelperegegeselqntaysnmn 180
QY 521 QPRVVELMRPEPSKSLGISIVGSGRMSRLSNGEVMRGIFIKHVLSDSPACKNGTLKPGDR 580
DB 181 qprvvelmrpepsksglgsivgsgrmsrlsngewmrgiflkhvledspakngtlkpgdr 240
QY 581 IVE 583
DB 241 lve 243

RESULT 14
AAV74151
ID AAV74151 standard; Protein: 206 AA.
XX
AC AAV74151;
XX
DT 14-MAR-2000 (first entry)
XX
DE Human prostate tumor EST fragment derived protein #338.
XX
XX Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;
KW treatment.
XX
XX Homo sapiens.
XX
XX DE19820190-A1.
XX
XX 04-NOV-1999.
XX
XX 28-APR-1998; 98DE-1020190.
XX

PR 28-APR-1998; 98DE-1020190.
 XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
 PA Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
 XX MPI; 1999-621386/54.
 DR N-PSDB; AA5252969.
 XX
 PT New human nucleic acid sequences from pancreatic tumors, and related
 PT proteins -
 XX
 PS Claim 23; Page 450; 502pp; German.
 CC
 CC This invention describes novel polypeptides and their encoding nucleic
 CC acids derived from human pancreatic tumor tissue which have cytostatic
 CC activity. The sequences are also useful in producing pharmaceutical
 CC compositions for treatment of pancreatic tumors. AA573814-Y74252
 CC represent protein fragments encoded by the human pancreatic tumor cDNA
 CC library derived expressed sequence tag (EST) sequences represented in
 CC AA52858-253014.
 XX
 SQ Sequence 206 AA:

Query Match 13.3%; Score 928; DB 20; Length 206;
 Best Local Similarity 95.3%; Pred. No. 6.5e-67;
 Matches 184; Conservative 1; Mismatches 6; Indels 2; Gaps 1;

OY 606 SAFAEKSDHTOSSAKTSQDVDEKDEFGYSWKNIERYGTLTGELHMLEKSHSGTGL 665
 DB 1 sfaeagschhtgsasaklsqdvdekedefyswknlererygtltgelhmtelaksgnsl 60
 OY 666 SLAKKNDKSRMSVFTVIGIDPNCAGKGRLOJADELEETNGIILYGRSHQNSSTIICAP 725
 DB 61 slagkhkdsrmvflvgldpngaagkdgrlqadelletngqillygrshqnsstikcap 120
 OY 726 SKVKIIFIRNKDAVNOMAVCPGNAVEPLPSNSENLONKTEPTVTSDAAVLLSFKNVQ 785
 DB 121 skvkiifirnkdaavnqavcpnaveplpsnsenlnqketepvtcsdaavllsfnvq 180
 OY 786 HLELPKDOGIGI 798
 DB 181 h--sgasgsggrgv 191

RESULT 15

AAV04736
 ID AAV04736 standard; Protein; 590 AA.

AAV04736;

06-JUL-1999 (first entry)

PDZ domain-containing protein gene encoded by clone D-2.

PDZ domain; gene expression: human umbilical vascular endothelial cell;
 HUVEC; stimulation; tumor necrosis factor; TNF; protein binding; PCR;
 cell; proliferation disorder; cancer; primer; amplification.

Homo sapiens.

MO9907846-A1.

18-FEB-1999.

12-AUG-1998; 98WO-JP03603.

19-JUN-1998; 98JP-0189944.

12-AUG-1997; 97JP-0230356.

(CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.

PI Funahashi S, Miyata S;
 XX MPI; 1999-167423/14.
 DR N-PSDB; AAX29967.
 XX
 PT Protein containing PDZ domain, whose expression is enhanced by TNF
 PT stimulation - plays an important role in protein/protein
 PT interactions and is used for screening for proteins for use in
 PT treatment of cell proliferation disorders such as cancer
 XX
 PS Example 7; Page 140-146; 240pp; Japanese.
 CC
 CC This sequence represents a new protein containing PDZ domains encoded
 CC by clone D-2, whose expression in human umbilical vascular endothelial
 CC cells (HUVEC) is enhanced by stimulation with tumor necrosis factor
 CC (TNF) alpha. The new protein is used to identify proteins which bind
 CC to it (particularly to the PDZ domains) and the genes encoding them,
 CC for use in the treatment of cell proliferation disorders such as cancer.
 XX
 SQ Sequence 590 AA:

Query Match 11.6%; Score 814; DB 20; Length 590;
 Best Local Similarity 98.8%; Pred. No. 6.6e-57;
 Matches 160; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MYCCRRTPPTTQSLDSDLCIDITLTKPHVDLGEFTGSSSETEDPVLAAMDAGSTEEV 60
 DB 429 myccrrtvppttqselddldidcltelkphvdlgeftgsssepepvlamtdagsteev 488
 OY 61 CAPLAMEAGIOHILEKSGKLGFSILDYODPIPDASTVITIRSLVPGIAEKGRLLP 120
 DB 489 qaplammeeagihmleksgkylgfsildyqdpdpastvilitrslvpglaekgrllp 548
 OY 121 GDRLMFVNDVNLNSLSLEAEVAKGAPSGTVRIGVAPRLP 162
 DB 549 gdrlmfvndvnlenslleaevalkgapsgtvrigvaprlp 590

Search completed: July 12, 2001, 14:40:26
 Job time: 189 sec

Fri Jul 13 15:00:06 2001

us-09-502-698-1.rag

Page 14

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2001, 14:38:32 ; Search time 40.62 seconds
(without alignments)
2574.782 Million cell updates/sec

Title: US-09-502-698-1

Perfect score: 6990

Sequence: 1 MCCCRRVPTTQSELDSD.....EENVAIIKRTKGVTLNVL 1373

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 7617452 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5811	83.1	2054	2	T46612 multi PDZ domain p
2	5617	80.4	2055	2	T30259 multiple PDZ domain
3	962.5	13.8	2172	2	T20145 hypothetical prote
4	589	8.4	728	2	T09457 numb-binding prote
5	586	8.4	628	2	T09458 numb-binding prote
6	534	7.6	1012	2	T23160 hypothetical prote
7	523	7.5	2450	2	S71625 protein-tyrosine-p
8	510.5	7.3	2294	2	T67630 protein tyrosine p
9	509.5	7.3	2466	2	I67629 protein tyrosine p
10	494.5	7.1	2490	1	A54971 protein-tyrosine-p
11	453.5	6.5	1256	2	JEO209 brain-specific ang
12	407.5	5.8	852	2	T10811 channel associated
13	399.5	5.7	1131	2	T15617 hypothetical prote
14	398.5	5.7	1277	2	T14152 synaptic scaffold
15	397.5	5.7	870	2	G01974 channel associated
16	385	5.5	1171	2	T42372 probable guanylate
17	384	5.5	767	2	T09599 postsynaptic densi
18	383	5.5	724	2	JH0800 synapse-associated
19	380	5.4	720	2	A45436 synapse-associated
20	379	5.4	960	1	A39651 discs-large tumor
21	378.5	5.4	911	2	I56552 synapse-associated
22	377	5.4	904	2	I38757 homolog of Drosoph
23	377	5.4	926	2	I38756 homolog of Drosoph
24	357	5.1	1112	2	T33733 AMPA glutamate rec
25	353	5.0	1464	2	T13711 bazoeka gene prote
26	352	5.0	1337	2	T13948 acylcyl protein k
27	264	3.8	1360	2	T34302 cell polarity prot
28	246	3.5	1163	2	JEO366 tight junction pro
29	243	3.5	1736	2	A47747 tight junction pro

30	237.5	3.4	1745	2	A46431 tight junction-ass
31	237	3.4	1495	2	T31434 desmin-180 - rat
32	233.5	3.3	1281	2	T00346 hypothetical prote
33	228	3.3	1116	2	I54378 gene x104 protein
34	216	3.1	87	2	S60315 PSD-95-related pro
35	214.5	3.1	1095	2	T43275 neurebin - rat
36	210	3.0	1367	2	T13703 tana protein - fru
37	203.5	2.9	2232	2	T34434 hypothetical prote
38	202	2.9	13055	2	T16580 hypothetical prote
39	197.5	2.8	3562	2	A47171 chondroitin sulfat
40	197	2.8	505	2	S62894 alpha-syntrophin -
41	194.5	2.8	126	2	I81210 tyrosine phosphata
42	194	2.8	1829	2	T24583 hypothetical prote
43	191	2.7	723	2	T14765 hypothetical prote
44	190.5	2.7	1829	2	T41751 1-atriadin - rat
45	190	2.7	817	2	T03852 protein phosphatas

ALIGNMENTS

RESULT 1

T46612 multi PDZ domain protein 1 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 18-Feb-2000 #sequence:revision 18-Feb-2000 #text_change 21-Jul-2000

C:Accession: T46612

R:Ullmer, C.; Schumack, K.; Figge, A.; Lubbert, H.

FEBS Lett. 424, 63-68, 1998

A:Title: Cloning and characterization of MUPP1, a novel PDZ domain protein.

A:Reference number: 223104; MUID:98196865

A:Accession: T46612

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-2054 <URL>

A:Cross-references: EMBL:AJ001320; NID:g2959978; PIDN:CAA04681.1; PID:g2959979

A:Experimental source: brain

C:Genetics:

A:Gene: MUPP1

Query Match	83.1%	Score 5811	DB 2	Length 2054
Best local similarity	79.5%	Pred. No. 0		
Matches 1147	Conservative 102	Mismatches 116	Indels 78	Gaps 5
QY	1	MCCCRRVPTTQSELDSDLDLDIELEKRPVLDGEFFISSETEDEPVLAMTDAGOSTEEV	60	
DB	620	MCCCRRVPTTQSELDSDLDLDIELEKRPVLDGEFFISSETEDEPVLAMTDAGOSTEEV	679	
QY	61	QAPLAWMEAGIQHIELEKSGKGFSLDYPIDPASTVYIIRSLVPGGIAEKDRLRP	120	
DB	680	QAPLAWMEAGIQHIELEKSGKGFSLDYPIDPASTVYIIRSLVPGGIAEKDRLRP	739	
QY	121	GRIPLFVNQVNLNENSSLEAEALKCAPSGTAVIGVAKPLPSPEEGYSAKEDSLTLP	180	
DB	740	GRIPLFVNQVNLNENSSLEAEALKCAPSGTAVIGVAKPLPSPEEGYSAKEDSLTLP	799	
QY	181	HSCEERGLADKPLFRADLALVGTNDALVDESTFESPPENDSYSTQASITLSHGSSC	240	
DB	800	HSCEERGLADKPLFRADLALVGTNDALVDESTFESPPENDSYSTQASITLSHGSSC	859	
QY	241	GGIINAGSSLPSPKDYENSCDPVLDLHMSLEELYTONLEROEDNPSVDISGAPS	300	
DB	860	GGIINAGSSLPSPKDYENSCDPVLDLHMSLEELYTONLEROEDNPSVDISGAPS	918	
QY	301	GFTINDYTPANAEQYECENTIVTESHLPSEVISAEPLVLPDSAGSGEHLDEQSS	360	
DB	919	GFTINDYTPANAEQYECENTIVTESHLPSEVISAEPLVLPDSAGSGEHLDEQSS	974	
QY	361	LACNAECVLMQNVSKESFEFTINIAKGNSSLGMTYSANKDGLGMIYRSIIHGSAISRDR	420	
DB	975	LVSDEASVTLQSKQAEAFERTVYIAKSSSLGTMYSANKDGLGMIYRSIIHGSAISRDR	1034	

QY	421	IABGCIJLISNEESTISTYNAOAMRRRLSLGPOKITYVAEHLFEFKSLGQOSR	480
Db	1035	IAVGCIJLISNEESTISTYNAOAMRRRLSLGPOKITYVAEHLFEFRVSGQAG	1094
QY	481	VALADIFSSYGRDIPELPEREBEGESESLQNTAYSMNMPRRVLEWREPSKLSISIV	540
Db	1095	IMALDIFSSYGRDIPELPEREBEGESESLQNAAYSMNQPRRVLEWREPSKLSISIV	1154
QY	541	GGRGMSLNSGEVWRGIFITKHVLEDSPACKNGTLKRGDIVE-----	583
Db	1155	GGRGMSRLNSGEVWRGIFITKHVLEDSPACKNGTLKRGDIVEVDCMDRDSASHEQAVEA	1214
QY	584	-----APSOSES	590
Db	1215	IRKAGSPYVFNVOSTIVNRPKRSPILSPHLKRCSSSTNPFAESLQLTSDKAPSOSES	1274
QY	591	EPKAPLCSVPPPPPSAFEMGSDHTQSSAKISODVDEKEDFEGYSWKNIIRERYGLTGE	650
Db	1275	ESEKRTLCSSVSSPSVSEKSSPOAPSATVVADEDEKDEEGYSWKNIIOERYGLTGO	1334
QY	651	LHMTLEKSHGGLSLAGNDRSRMSVFTVGDIPDNAGKORLOIADLELEINQILY	710
Db	1335	LHMTLEKSHGGLSLAGNDRSRMSVFTVGDIPDNAGRDRLOIADLELEINQILY	1394
QY	711	GRSHONASSIIKCAPSKYKIIIFIRKNDAVNQMAVCPGNAEPLSPNSENLQKETEPTVT	770
Db	1395	GRSHONASSIIKCAPSKYKIIIFIRKNADAVNQMAVCPGSAADPLSPNSEPOKKEVEPST	1454
QY	771	TSDAVNDLSSKRNQHLLEPRDQGGGLAISSEDTLSGYIKSLTSHGYAALDGRKVD	830
Db	1455	TSASAVNDLSSLTNYVHELPRDQGGGLAICEEDTLNGVYIKSLTSHRGGAANDGRKLPD	1514
QY	831	QILAVDDEIVVGYPIEKFISILTKAKTKVTLTTHAENPDQAPASAGAASEKKRSSOS	890
Db	1515	RIAVDDELVAHGCIIEKFIISLTKAKTKVTLTGAENPGQAPASAAVYASERKDSQST	1574
QY	891	LMVPOSGPPEPESIRNTRSRSTPAIRFASDPATCPIIPGCTTIEISKRTGJGLSIVGSS	950
Db	1575	PAPV---APDLEPPISTSRSTPAIRFASDPATCPIIPGCTTIEISKRGOTGJGLSIVGSS	1633
QY	951	DTLLGAFIIEHYEBEGAACKDGRLMADQJILEVNGIDLRKATHEAIVNLRQTPORVRLT	1014
Db	1632	DTLLGAIIEHYEBEGAACKDGRLMADQJILEVNGIDLRKATHEAIVNLRQTPORVRLT	1694
QY	1011	LYRDEAPYKEEBEVDCTLTIELOKKPGKGLSIVYKRNQTVFSDIVYKGIADPDGRIL	1070
Db	1692	LYRDEAPYKEEBEVDCTLTIELOKKPGKGLSIVYKRNQTVFSDIVYKGIADPDGRIL	1751
QY	1071	OGDOILLVNGSDVNAOSEVAALIKCSLGTIVLLENGRIKAGFSEHERPSQTSOYSEGS	1133
Db	1752	OGDOILLVNGSDVNAOSEVAALIKCSLGTIVLLENGRIKAAFFSEHERPSQSOSYSEGS	1811
QY	1131	LSSFPTPLSSGSSTSESSLESSKKNALASEIQGRJTYEMKKGPPDSDGISIAGVGSPLGD	1190
Db	1812	LSSFPLPSRSGIHSTSESSSSAKKNALASEIQGRJTYEIKKGPDALGLSIAGVGSPLGD	1874
QY	1191	VPIFIAMMHPRPGVAAOQOKLAVGDRYITTCGSTSTGHTHQAANLKNKNSGSIEVOYVAG	1254
Db	1872	VPIFIAMMHPRPGVAAOQOKLAVGDRYITTCGSTSTGHTHQAANLKNKNSGSIEVOYVAG	1933
QY	1251	GDVSVVTHGHOEPASSLSIFTLGSLSTIFPDGDPPOCKSTILERPDGAGFSIYGVGYS	1311
Db	1932	GDVSVVTHGHOELANPCLAFTGLTSITTFPDGDPPOCKSTITILDRKPDGLGFSIYGVGYS	1991
QY	1311	PHGDLPIYKTVFAKGAASEDGRKRGDOJIIAVNGOSLEGVTHHEAVAIKRTKGTVTLM	1370
Db	1992	PHGDLPIYKTVFAKGAASEDGRKRGDOJIIAVNGOSLEGVTHHEAVAIKRTKGTVTLM	2051
QY	1371	VLS 1373	
Db	2052	VLS 2054	

```

RESULT      2
130259 multiple PDZ domain protein - mouse
C:Species: Mus musculus (house mouse)
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000
C:Accession: T30259
R:Simpson, E.H.; Suffolk, R.; Jackson, I.J.
Genomics 59, 102-104, 1999
A:title: Identification, sequence, and mapping of mouse multiple PDZ domain protein g
A:Reference number: z20797; MUID:99326529
A:Accession: T30259
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-205 <SIM>
A:Cross-references: EMBL:AJ131869, NID:g4150877, PIDN:CAAI0533.1, PID:g4150878
A:Experimental source: strain C57/Bl6 x CBA Fl; whole brain
C:Genetics:
A:Gene: mpdz
A:Map position: 4

Query Match          80.4%; Score 5617; DB 2; Length 2055;
Best Local Similarity 77.1%; Pred. No. 0;
Matches 1112; Conservative 114; Mismatches 139; Indels 78; Caps 5;

QY   1 MVCRCRTVPTTQSGLSDLDLCIDLETKEPHNDGEGFISSETEDPVLAMDAGOSTEEV 60
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   621 MVCRCRTVPPIALISEDSIDLINDELTERPHDLGEFGSSETEDPMLMSDVQNMEEI 680

QY   61 QAPLLMMEAGIOHIELEKSGSKGLFSILDYODPDIPDASTVLIIRSLVGGIAEKDGRLP 120
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   681 QTPIPLAMMEAGGGSIELEKSGRGIFSIILDYODPIDPATVTLYIRSLVGGIAEKDGRLF 740

QY   121 GGRLMFVDNVNIENSSLEEVAVALKGAPSGYIRIGAVAKPLPSPEEYGVSAAKDESFLLP 180
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   741 GGRLMFVDNINENSTLEEVAVALKGAPSGMYRIGVAKPLPSPEEYGVSAAKDEAFICSP 800

QY   181 HSCCEAGLADKRPLFRADIALVTGNDAIDVEESTFEPSPPENDSYSTQAASITLSLHGSSC 240
    ||:::||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   801 HACKEESGLDKALFRADIALIDTPAEESIABREFSQSPDNDSYSTQAASITFLSHGTC 860

QY   241 GDGLATGSSLPSPKDYIENSCLPYLDLHMSLEELYQNILLERODENTPSVDISMGPAS 300
    ||||| | | | | | | | | : | | | | | | | | | | | | | | | | | | | | | |
Db   861 SGGMWNGPFLSSSPKDYVSS--EVYLGIHLISLEELTYQNILLEROHASSSTDMKPPTPF 919

QY   301 GFTINDYPRANILEOYECEMTIWTESHLPSEVISSEALPSYLDGSAKGSEHLEQSS 360
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   920 GFXTSDYTNTNAGECKIQOMANPAPHSQLPSN-LSTELASAL---AAVAPKTLTDQTF 975

QY   361 LACNAECVAILQVSKESPERTINIKAGNSSLGMTYSANKDKGLGVRSIIINGGAISRDR 420
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   976 LASDAESVTLQSMQEAEPFYLTIAKSSSLGMTYSANKDKGLGYIVRSIIINGGAISRGR 1035

QY   421 IAIQDCILISINEESTIVTNAQARAMLRRHSILGDIKITTVPAHEHEEFKISLGOOSGR 480
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   1036 IAYGCILISINEESTISLTNAQARAMLRHSILGDIKITTVPAHEHEEFKISLGOQAGG 1095

QY   481 VVALDFSSYTGRODPELPEREGESESELONTAYSMMNOPRAYELMRREBSKISGISIV 540
    ::|||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   1096 IVALDFSSYTGRODPELPEREGESESELONAAVSSWSOPRAYELMRREBSKISGISIV 1155

QY   541 GGRGMSGRLSNENVRKGFIFKHVLEDSPAGKNGTALKPGDRIVE----- 583
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   1156 GGRGMSGRLSNENVRKGFIFKHVLEDSPAGKNGTALKPGDRILIIEVDGMGLRASHDAQVEA 1215

QY   584 -----APSQSES 590
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   1216 IRKAQNPVFVFWQSIINRPKSPPLSLPHSLYPKYSEFSTNPFADSLQLTTDAQPSQSES 1275

QY   591 EEKKAPLCSVPPPPPSARAEMGSDHTQSSASKISQDVVKEDPEFGSMKNIIEERGTLTGE 650
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   1276 EEKAPLCSVPPPPSPSVSEMGSDCAQPSATVAVSDEKEDPEFGSMKNIIDERRGSLTQG 1335

```

[illegible]

A: Molecule type: DNA
A: Residues: 1-2172 <WT>
A: Cross-references: EMBL:Z46795, PDB:1CA86789.1, GSPDB:GN00020, CESP:C52A11.4
A: Experimental source: clone T19E10
C: Genomics:
A: Gene: CESP:C52A11.4
A: Map position: 2
A: Introns: 27/2; 78/1, 111/3, 156/2, 208/1, 230/3, 290/2, 341/1, 389/3, 431/1, 461/3, 24/3, 1747/1, 1946/3, 2049/3, 2096/3, 2145/3

A:Accession: T09457
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A.Residues: 1-728 <DB>
A:Cross-references EMBL:AF034745; NID:g3041878; PIDN:AAC40075.1; PID:g3041879

C:Genetics:
C:Experimental source: developmental stage: embryo
A:Gene: lnx
C:Superfamily: RING finger homology
F:41-88/Domain: RING finger homology <RNR>

Query Match 8.4%; Score 589; DB 2; Length 728;
Best Local Similarity 29.1%; Pred. No. 4,7e+25;
Matches 188; Conservative 115; Mismatches 236; Indels 106; Gaps 23:

Dy 779 SSKFNVOHLELPKD-----QGL-GIAISEEDTLGVIIKSLTEHGVAATDGRUKVGDO 831
| : | : | : | : | : | : | : | : | : | : |
Db 130 TSCKASHYGLTKDKRRSODCCPDGCASIMATTLS-----PEVSAAT----- 173

Oy 832 ILAVDELIVGPIKEFTSLKTAKMTVKLTITHAENPSONAP---SAAAGASCERKNSS 888
| : | : | : | : | : | : | : | : | : | : |
Db 174 ISLMDEPGLDNP--AYVSVEDGEPAVNSSCGSNRTTRAPERSTWRSRSEFKTINA 231

Oy 889 QS-LMWPOSGSPESPT----NTRSSTPALFASDPATCPRIIPCE-TTIEISKRGTL 942
| : | : | : | : | : | : | : | : | : | : |
Db 232 LIALRTKSGSVANHHVDQRDNSENTTYPEVF---PRFLHLDPGETISIKINRADPSE 288

Oy 943 GLSI--VEGSDPLLGALFIHEHYEEGAACKORLAGOOILEVNGCIDRKATHDAIYNVL 1000
| : | : | : | : | : | : | : | : | : | : |
Db 289 SLISILVGGSERPLEHIITIOTIRGDVTARDRLRPGLDIILVANGMDISNNPHNYARLL 348

Oy 1001 ROTPORVRILTYYRDE-----APYEKEYCDTLTIELOK-KPKKGIGLSIYGKRNDY 1050
| : | : | : | : | : | : | : | : | : | : |
Db 349 ROPCOQLRTLTVLRREQKFRRSRNAHVDPDSYGPREDPSFHVIANKSSPEEDQGKITVARVPDR 408

Oy 1051 GFVFSDIYKKGIADDDGDLIQDGDTLLVNGEDEVRYNASOEBAVALLKCSLYTYLEVGNK 1110
| : | : | : | : | : | : | : | : | : | : |
Db 409 GFETINVNLNGVADRHRGOELEENDRVLAINHDLRFGPSBSAHAILQAOSERRVHALVSR-- 466

Oy 1111 AGPFHSERPSP-----QTQSVCSEGSLSFTHPLGSSTSSESLESSKKNALASEIOGLRT 1165
| : | : | : | : | : | : | : | : | : | : |
Db 467 -----QWASSPDIREQAGWNINSQQS-----PGPER-----NTASKRAANICHE----KV 508

Oy 1166 VEMKKGPDPDSLGISIAGVGSPLGDVPPIFIAMHPGTVAATOCTKLAVGDRVITYICGSTE 1225
| : | : | : | : | : | : | : | : | : | : |
Db 509 VSVMKPDESISLGMTTGFGASHREMDLPITYVISVEEGGYISRGRITKTDIDLNVNGIELT 568

Oy 1226 GMTHTOAANLLKNMAGSTEMOVVAGGDVVYTGHNOEPASSLSFTGLTSTSIFPDOLDGP 1285
| : | : | : | : | : | : | : | : | : | : |
Db 569 EVSRTEANAILLKAPSASYLVKANL---EVK-----EQEAODECSFALLDS-----NHANTP 615

Oy 1286 P-----SQSKITLEGRPDC-IQFSIVGVGSFPHDGPLTYKTVFAKG 1326
| : | : | : | : | : | : | : | : | : | : |
Db 616 PGDMSPSWMMLELTPOLYLNCNKDVILRNRTAGSLGFCAIVGGEYSGNKPPFIKIVEGT 675

Oy 1327 AASEBGRALKRDQIIIAVNGSOLECYTHEENAAILKRTGCTWTMLMS 1373
| : | : | : | : | : | : | : | : | : | : |
Db 676 PAYNDGRTRCDDILLAVNGRSTSGMIHACLARMKELGRITLTITAS 722

RESULT 5
T09458
numb-binding protein LNXp70 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T09458
R:Dho, S.E.; Jacob, S.; Wolting, C.D.; French, M.B.; Rohrschneider, L.R.; McGlade, C.
J. Biol. Chem. 273, 9179-9187, 1998
A>Title: The mammalian numb phosphotyrosine binding domain: Characterization of bindin
A:Reference number: Z16678; MUID:98204916
A:Status: preliminary; translated from GB/EMBL/DDBJ

[illegible]

A:Molecule type: mRNA
 A:Residues: 1-628 <DHO>
 A:Cross-references: EMBL:AF034746; NID:93041880; PIDN:ANC40076.1; PID:93041881
 A:Experimental source: brain
 C:Genetics:
 A:Gene: lnx

Query Match 8 4%; Score 586; DB 2; Length 628;
 Best Local Similarity 29.0%; Pred.No.5,4e-25;
 Matches 187; Conservative 114; Mismatches 237; Indels 106; Gaps 23;

QY 782 KNYQHLEPKD-----QGL- -GIAISEDTLSGVITKSLTEHGVAATDGRKVGQIILA 834
 DB 33 KASHGGLTKDKRRSQDQDCPCGASLMATTL-----PEVSAAT-----ISL 76
 QY 835 VDDEIVVGVPIEKISLTKAKTKVKTTHAENPDSQAVP---SAGAAAGEKKNSSQS- 890
 DB 77 MTEDEPLDNP--AYSSVEGEPEVANSSDGSRNTRARPERSTWRSRSEFKKINRALS 134
 QY 891 LMPVQSGSEPESEI-----KNTSRSTPAIFASDPATCPIIPGCE-TTIEISKRTGLGIS 945
 DB 135 LRRTKSGSVANVNHVDGQDNSENTEVTVEVF--PRLFHILPDGEITSIKINRADPSESLS 191
 QY 946 I--VGSDFLLGAFIIEHYEGGACCKDGLMAGDQILEVNGIDLRKATHEAIVNLRQT 1003
 DB 192 IRLVGSSETPVNIITIIQHTYRDGVIAKDRGLRGLDIIILKNGMDISNPHNVAIVLLRQD 251
 QY 1004 PORVRLTLRYDE-----APYKEEVEODTLTIELQK-KPGKGLSLSVGRNDTVF 1053
 DB 252 CVLRLTLVREQKFRSRNANHVPDSYCPRODSFHVILNKSSPEQGIKLVVRVDEPGVF 311
 QY 1054 VSDIYVGGIADPPGRILQGDQIILLVNGEDYRNASQEAVALKLSGLTYLVEGRKAC 1113
 DB 312 IFNVLLGVAAHRGQLLENDRLAINGHDLREFSPESAHLIQASERRHVLVSR----- 366
 QY 1114 FHSERRPS-----QTSQVBSGLSSTFPLSGSSSTESSKKNALASEIGLRTVEK 1168
 DB 367 --QVRRSSPDITQEGAMISNGQS-----PQGR-----NTASKPRATCHE---KVYSV 411
 QY 1169 KKGPTDSLGISIAGVSGSPGLDVPIFIAMHPTGVAAGTQKLRVGRIVTICSTEGMT 1228
 DB 412 WKPPSSSLGWTVGCGASHREMDPIYIVISVEPGVISRDRKITGDIILVNGIELEVS 471
 QY 1229 HNGAVLLKNASISIMOVYAGDGVYVGHNEBPASSSLSTFGLTSTSFQDLDLP-- 1286
 DB 472 RTEAVAILKASPSVVLKAL--EVK-----EQEAQEDCSPALDS---NNHVTTPGD 518
 QY 1287 -----QKSTLERGPDG-LGFSIVGVGSPHGDLPYIKVTFAGKAAS 1329
 DB 519 WPSWVWMLLEPOLYLCNCKDVIILRNRTAGSLGCIYGVGEIYSGNRPFFIKSIVEGPAT 578
 QY 1330 EDGRKRGDQIIVANGOSLEGVTHEEVAVALKRTKGTVTLMVL 1373
 DB 579 NDGRICGDILLAVNGRSTSGMIHACLARMLKELKGRITLTIAS 622

RESULT 6
 T23160
 hypothetical protein K01A6.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15 -Oct-1999 #text_change 18-Aug-2000
 C:Accession: T23160
 R:Cottage, A.
 submitted to the EMBL Data Library, January 1996
 A:Reference number: Z19701
 A:Accession: T23160
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1012 <ML>
 A:Cross-references: EMBL:268750; PIDN:CAA92965.1; GSPDB:GN00022; CESP:K01A6.1
 A:Experimental source: clone K01A6
 C:Genetics:

A:Gene: CESP:K01A6.1
 A:Map position: 4
 A:Introns: 107/1; 151/2; 183/2; 270/2; 316/3; 410/3; 628/3; 922/3; 971/3.
 C:Superfamily: WW repeat homology
 F:131-168/Domain: WW repeat homology <WWR>

Query Match 7 6%; Score 534; DB 2; Length 1012;
 Best Local Similarity 24.1%; Pred.No.9,6e-22;
 Matches 208; Conservative 123; Mismatches 298; Indels 234; Gaps 31;

QY 572 NGTLKPGDRIEAPSSQSESEPEKAPLCVSP-----PPPAFAFAMGSDHTQSSAKIS 624
 DB 317 SGTLK-----SSSPRDSGFDSSPTTRKRGDPERATSSADVDHSHKMSRSS 365
 QY 625 QDVKDEDFGYSKKNIRERYGTLGELHNIIELEKHSGLGLSLAGNKRDSRVFT--VG 682
 DB 366 NPLFTTDP-----ARLGELISTKIKVKGAKGLGFTLIGNDSSRDEFTQVKS 413
 QY 683 IDPNGAAGDKRLQIDELLEINGQILYGRSHONASSIICKAPSKYKIIIFINKDAVNO- 741
 DB 414 VLSGPGAAANGVLRSGDILLVRNGRLLLGATQKEACDVEVAIP-----VNDA 460
 QY 742 --MAVCPGNAVEPLPENSENLQKETEPTVTTSDAAVDSFKKNVOHLEPKDQGLGIA 799
 DB 461 VDIQVGRGELFDIPAN-----RIVTENYAAKS-RDLHEIDIFKSGSEGFPT 508
 QY 800 ISEEDTLGVITKSLTEHGVAAATDGRKLVQDQILAVDEIVVGVPIEKISLTKAKTV 859
 DB 509 IA--DNLNGORIKIL---FPSQCPMLMEGDTIVELDEGNVRPIPTQLVDMLREPIGY 563
 QY 860 --KLTHAENPDSQAVPSAAGAAGEKKNSSQSLMP-----QSGSPESIR 905
 DB 564 RGLLVKRSKPTKRSPTPAARFYRGPTNMDSAAPLVRSKTPAERQTSREEDQNV 623
 QY 906 NTSRSTPAIFASDPATCPIIPCEFTTIEISKRTGLGISVGSDFLLGAFIIEHYEE 965
 DB 624 NT-----LQRPVATSEWEMSSAIPASRRP-----STTLTG- 657
 QY 966 GAACKDRLMAGDQILEVNGIDLRKATHEAIVNLRQTPORVRLTYRDEAPYKEEVC 1025
 DB 658 -----ATPNYIPLSQ-----NQKPSD 674
 QY 1026 TLTIELQKPGKGLSIY--GRNDTGVSVDIVKGIADPPGRILQGDQIILLVNGEDR 1084
 DB 675 LITVSLIRKP-VGEGRLLGVESKTPLSVGQIVIGAAEEDGRLOEGEIVEIDSHNV 733
 QY 1085 NASQEAVALKLSGLCT--VTLEVGRKAKPFHSERRPQTSQVSESSLSFTPLSGSS 1142
 DB 734 GASHSEAVVLEAAAKNHVKLY-----RRPRTDPARGSLNS-----AGPS 777
 QY 1143 TSELESSSKKNALASEIOGLRTVEKKGPTDLSISINGV---GSLGDNPIFIAMH 1199
 DB 778 -----GSYDVLLHRRNDEGFGVLLMSQHKNSSTVGQI-----Q 811
 QY 1200 PTGVAOTKLRVGRIVTICSTEGMTHTQAVNLLKNASISIMQV---VAGDVS 1255
 DB 812 PGSPAAKGRSLVGDRIAVANGIDILSLSPDTISLKDSGLSVRLTINPPTAGVLP 871
 QY 1256 VT-----GHHO-----EPASSLSFTGLTSTSFQDLDG--PPQK 1289
 DB 872 VSAITLGRNFTMNGHYESNGLPPPPSYVEKHPPEYLAIFDLS-----INDRSMNGNLI 927
 QY 1290 SITLERGPDGLGSIYGVGSPHGDLPYIKVTFAGKAASEGRKRGDQIIVANGOSLE 1349
 DB 928 DVTLEGTGFGFSING--GOEFGSMPLFVLRITADGPAKADRLQVGDQLTTINGQSTK 985
 QY 1350 GVTHEEVAVALKRTKGTVTLMVL 1372
 DB 986 GMSHDAIRIIRKO-HTMWNLTVL 1007

RESULT 7


```

Db      326 PEAVYVTRSTTRKKARISDGSIALDIFGPKMDPIYTRRELPTS---SALSSALDRIR 382
      268 DLHMSLEELYTONLERODENTPVSIDSMGASGFTINDYTPANAIEQOYEGENTIVMTE 327
      383 ERQKRL-----QVLRAMNVEEP-----VRKTTYHGDVFSRS 415
      328 SHLPSEVISAEPLPSVLPDSACK-----GSEHLEKQSSL-----ACN 364
      416 SSSPSIISSEDSFQVRRSEASKRFESSGGLPVDLTLGOGQSQRSEYETPEEGNLIN 475
      365 AECVMLOANSKESFERTINIAGNSSL-----GWTVASN-----KQDGLMIV-----406
      476 QE-IMLRQEBELMQLQAKMALRQSLSLPGDTIKASMLDITRDLRELALETANTQK 534
      407 -----RSIIHGSAISRDRGRIAGDCILSINESTISVTNAO- 442
      535 LRFNFGPEFVKMTIEPFIISDLPRSLITFKKKNEDNRKYN--IMLLNGQRLLETCDTKT 592
      443 -----ARAMLRRHSL-----IGPDITITYPAHLEE-----469
      593 ICKDVEPMVANHIGLVEHHLFALATLKDNEYFEVDPDLKLTVPARGMKEEPKKTKATV 652
      470 -----FKI-----SLGQSS-----478
      653 NTLFRIRKFMDDVSLIQLHTLCHQYIQLRKDIEERHNCDDETSLLASIALQAEYG 712
      479 -----GRVA-ALDIFSSYTGRIPELPEREEGE-ESLQNTAVS 517
      713 DYQPEVHGVSYFMEHYHLPARVMEKLDL-SYIKELPKLHNTYVASEKETELFEFLKYC 770
      518 NMNQPRVELMR-EPK-----SLGISIV-----GGRMGSRSLSGEVMRGIIFK 561
      771 QRLTERGVHHRVHPKKSOTGILLGVCSKGVLFVFNHNGVRLVLRFPRREKKSISFSK 830
      562 H--VLEDSFAG--KNGTLKRGDRIVEAPSOSESEPERAPLCVPPPPSAFEMGSDHTQS 618
      831 KKTTLQNTSGIKHGFQOTMSKICQYLLHNCSTQHFKQL-----QM 871
      619 SASKIQDVKDEDEFGYSKKNIR--ERYGTLTLELMIELEK-GHSGGLSLSLAGKDRSR 675
      872 RAKQSNQDODIDVLHRRMSIYSPER-----EITLVNKKAKKAGLGIOIGGKMRG 925
      676 --MSVTVIGDIPGAGKCDRLADLELEINCOILYGRSHOVASSIKCAPSKVITFI 733
      926 LDGRTISSVAPGCPADLDGCLKPGDRLLSVNSVSLGESHMAIETLQNAPEDEVILVIS 965
      734 RNKDAVQNAVCPGNAVEPLPSNSENLONKETEPTVTTSDAVALDLSFKN-----VOHL 787
      966 QPEKISKVSPSTVH-----LTNEMKNVMKKS-----YMQDSALDSSSKDHMSRGTLRHI 1037
      788 ELPK--DQGL--GIAISEEDTLSSVYIISLTHGVAALDGRKLVGDQDILAUNDEIVGY 843
      1038 SENSFSPSGLRGSLSSQDRTESASLSQSQVNGFFAS---HLDDQTV---QEOSHG 1090
      844 PIKFIISLKTAKMTVKTLTJHAENPDSQAVPSAAGASEKK-----NSOSLW 893
      1091 PPSVSVS-----KATEKETFTDSNOSKTKKPGISDVTVSDRSDSDMDEATVSSSDHOT 1145
      894 POGSPEPESIRNTSRSSPAIPASDPATCPRIIPGCTTIEISKRTGLSLSTVSGSDTL 953
      1146 PKQ-----ESSSSVNTSNKMNKFTSSSPK-----PGDIFVELAKKNNSIGISVTGCVNTS 1198
      954 L--GATIIHEVBEAGACADKRLMAGDQILEVNGIDIRKATHDDEALINVLKQTPQVRRLT 1011
      1199 VRHGIGYVKAIVIPQGAESDGRHKGDRVLAVNGVSLGATHRQAVETLINTQOVYHL 1258
      1012 YRDEAVYKEEV-----C-----DTLTIELKKPG 1036
      1259 EKQOSTSKHEVYVTPQCTISDNQAGQGEKAKTTQVADYFVTEENTFEVKLRKN-S 1317
      1037 KGLGLS-----YVGRNDTGVFVSDIYVKGGLADPDGRLIOGDOILLVNGEDVRNASOE 1089

```

```

Db      1318 SGLGFSFREDNLPIEQINASTVRYKKLEPGQPAESKIDVGDVILKNGASLKLGLSQ 1377
      1090 AVAALLKCSLGTVTLEVGRIKAG-----PFRSE-----1117
      1378 EYISALRGTAPEVFLLCRPPEGLVLEIDTALLTFLQSPAVQLPNSKDSQSPCEQST 1437
      1118 -----RRPQTSQVSEGS-----LSSTF-----1136
      1438 SSDENEMDSKSKKCKSPSRSDYSTDSGSGEDDLVTAPANISNSTWSALHQTLSNWS 1497
      1137 -----PLSGSST-----SESLESS 1150
      1498 QAOSHHEAKSDEDITCTMFYYPQKIPNKPEREDSNPPLPDMAFGOSYOPSSAS 1557
      1151 SK-----KNALASITQGLR-----TVEMKGPDSIGISLAG-- 1183
      1558 SMDKYHHIISEPTROENTPLKNDLENLEPELEVELLITLILIKSEKSLGFTVYKQ 1617
      1184 -VGSPLGDYPIFLAMHPGVAAGQOKLAVGDRIVTIGSTEGMTHQAVMLKNAGS 1242
      1618 RIGCYVHDV-----IQDP--AKSDGRLPGRDLIKVNDTVDVTNMTHTDAVNLRAASKT 1669
      1243 IEMQVYAGDVSVYTGHNQ--PASSLSFTGLTSTISFODDLGPPCKSITLERGPDGL 1301
      1670 VRL-----VIGRVLELRIPRL-----PHLLPDTITLCKNEELG 1703
      1302 FSTVGYGSPHGDPLIYKTVFAKGAASEDGRKRGDQIIAVNGSLSECTHBEA-----1356
      1704 FSLCGGHDSLY--QVYIYSDINPRVAAIEGNLDLDVHYVNGVSTQCMTELEVNRALD 1761
      1357 ---VALKRTGTVTLM 1370
      1762 MSPLSLVLKATRNDELTV 1779

RESULT 9
167629
protein tyrosine phosphatase (PTP-BAS, type 2) - human
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 16-Jul-1999
C:Accession: I67629
R:Maekawa, K.; Imagawa, N.; Nagamatsu, M.; Harada, S.
FEBS Lett. 337, 200-206, 1994
A:Title: Molecular cloning of a novel protein-tyrosine phosphatase containing a membr
A:Reference number: I53483; MUID:94116679
A:Accession: I67629
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2466 <RES>
A:Cross-references: GB:D21210; NID:9452191; PIDN:BA04751.1; PID:9452192
F:574-866/Domain: protein-tyrosine phosphatase, nonreceptor type 13; GLGF domain homolog
F:1354-1430/Domain: GLGF domain homology <GLG2>
F:2218-2437/Domain: protein-tyrosine phosphatase homology <PTP>

Query Match 7.3%; Score 509.5; DB 2; Length 2466;
Best Local Similarity 22.3%; Pred. No. 9, 8e-20;
Matches 256; Conservative 146; Mismatches 395; Indels 351; Gaps 41;

      510 ELQNTAVSMNQPRVRLMRREPESKSLGISIVGGRGMSRLSNG-----552
      868 QLOMRARQSNQADIERASFRLNQAESYRGFNNGRAISTGSLASLTINKLAVRPLSV 927
      553 --EVMRGIFIKHY-----LEDSPAGKNGTL---KPGD--RIEVARQSESEPERAPLCS 599
      928 QAEILKRLCSSELSTLOPLONSSKEKNDKASMEKPREMSKSYHDSQASLYPHRRNVIV 987
      600 VPPPPSAFAEM-GSDHTQSSAS-----KTSQDVDRKD 631
      988 NMEPPQTVAEIVLGKPSHMSRSDAESLAGVYTKLNSKSVASVLSNRPERRKHESDSSSIE 1047
      632 EFG--YSWKNIRERYGLNG---ELHMIIELEK-GHSGGLSLAGKDRSR--MSVTVGI 683

```

```

Db 1048 DPGQAVYLDVLRKMSIYSSPEREITLVNKKDAKYGLGFOITIGCEKMGRLDGIFFISSV 1107
Oy 684 DPNGAAGDGRLOIADLELLEINGOLYGRSHONASSIIKCAPSKYIIIFIRNKDAVNOMA 743
Db 1108 AFGPADLDGCLCPGDRILISVNSVSEGVSHAALEITLQNNAPEDVTLVISQPEKIKSV 1167
Oy 744 VCPGNAVEPLPSNSENLOKETEPTVTTSDAVIDLSFEKN-----VOHLELPK--DOGG 795
Db 1168 SPPVH-----LTNEMKNYKKS-----YMODSALDSSSDHMSRGTLNHISNNSGPGG 1219
Oy 796 L--GAISEEDTLGVIITKSLTEHGVAAITDGLKYGDOILAVDEIVVGYPIEFKISLTK 853
Db 1220 LREGSLSSQDSRTESASISQSOVNGFFAS-----HLGDQW---QESHGSPSPVIS--- 1269
Oy 854 TAKMIVKTLIHAENPDQAVPSAAGASGEKK-----NSSQSLMPPSSGPEDES 903
Db 1270 --KATEKETFDSDNSQSKKKKKEISDVITYSDRGSDMDENATYSSQDHQHPKQ--ESSS 1324
Oy 904 IRTNSRSSTPAIFASDPATCPRIIPCCETITELSKRTGLGISVGSPTLL--GAFIIE 961
Db 1325 SVNTSNKMNKFTFESSSPK-----PEDIFEVELAKNDNSLGISVGTVMVSRHGGIYKA 1380
Oy 962 VYEEBAACKDGRMLAGDOILEVNGIDELKATHEAINTLROTPOVRYRLTLRDEAPYKEE 1021
Db 1381 VIPGAAESDGRIRHKGDRVLAVNGVLEGATHKAKAVETLRMTGQVHLLLEKGSPTSKE 1440
Oy 1022 EV-----C-----DPLTLEOKKPKGGLGIS----- 1042
Db 1441 HVPYTPQCTLSDONAGGQPEKVKTTTQYKQSYTEENTPEVKLFKN--SSGLGSESRE 1499
Oy 1043 ---IYKRNQDGVPSVDIVKGIADPRDLTGDOILLVNGEDVYNASQEAVALLKCSL 1099
Db 1500 DNLPEQINASTIVRYKKLFEPGPAAESGKIDYGVITLKVNGASLKLGSQGEVTSALRGT 1559
Oy 1100 GTVILEVRIRAG-----PFHSE----- 1117
Db 1560 PEVELLILCRPPGVLPETIDTALLPLQSPAQVLPNNSKSSQSCVEOSTSDENEMSDK 1619
Oy 1118 ---RRPSQTSQVSEGS-----LSSFTF-----PL 1138
Db 1620 SKKOCKSFSRSDSYSDSSGSEGDLYTAPANISNTWSALHQLLSNVSQAQSHHEAPK 1679
Oy 1139 SGGST-----SESSLESSK----- 1152
Db 1680 SQEDTICTMYFPQKIPKPFEDSNPSLRPPDMAQSGSYQPSASSSSMDKXIHNI 1739
Oy 1153 -----KNALASEIOGLR-----IYEMKKGPTDSIGISTAG--VGSPLGVP 1192
Db 1740 SEPTROENWPLKNDLENHLEDFEVELLITLLIKSEKSGLFVYTKGNORIGCTVHGV- 1798
Oy 1193 IFIAMHPTGVAAGTOKLRVGDRIYITIGTSTEGTHTQAVNLKLNAGSITEMOVVAGCD 1252
Db 1799 ----IQDP-----AKSDGRKLKPDRLIKVNDTVDVNTHTDANLRAASKYVRL----- 1844
Oy 1253 VSVVTGHQHE-PASSSLSFTGLTSTSTIFQDGLGPPCKSTLERKPDGLGFSIVGYGSP 1311
Db 1845 ---VIGRAVLELPRIPLM-----PHLLPDIITLTCNKEKELGFSLGGHSL 1885
Oy 1312 HGDLPYVKTVPFAGGAASEDGRKRGDOIIIVNGOSLEGYVHEEA-----VALIKR 1362
Db 1886 Y--OVVYISDINPRSVAAIEGNLQDLIDVIRYNGVSTOGMTLEEVNRALDMSLSVLKA 1943
Oy 1363 TKGTVTLL 1370
Db 1944 TRNDLPV 1951

```

RESULT 10
A54971
protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type 13 - human
N:Alternate names: Fas-associated phosphatase FAP-1; protein-tyrosine-phosphatase hPTP1B
C:Species: Homo sapiens (man)

C:Date: 11-Nov-1994 #sequence,revision 08-Feb-1996 #text-change 01-Dec-2000
C:Accession: A54971; A55114; I59595; I53483; S46955
R:Barville, D.; Ahmad, S.; Stocco, R.; Shen, S.H.
J. Biol. Chem. 269, 22320-22327, 1994
A:Title: A novel protein-tyrosine phosphatase with homology to both the cytoskeletal
A:Reference number: A54971; MUID:94350988
A:Accession: A54971
A:Molecule type: mRNA
A:Residues: 1-2490 <BAN>
A:Cross-references: GB:U2128
A>Note: Sequence shown follows authors' translation at positions 62-63
J:Satas, J.; Claesson-Welsh, L.; Heldin, C.H.; Gopez, L.J.
J. Biol. Chem. 269, 24082-24089, 1994
A:Title: Cloning and characterization of PTP1, a protein tyrosine phosphatase with s
A:Reference number: A55114; MUID:95014139
A:Accession: A55114
A:Molecule type: mRNA
A:Residues: 1-61, 'GS', 64-839, 'D', 841-1055, 1075-1133, 'PH', 1136-1210, 'T', 1212-1383, 1389
A:Cross-references: GB:X80289; NID:9515030; PIDN:CMA56563.1; PID:9515031
R:Sato, T.; Irie, S.; Kitada, S.; Reed, J.C.
Science 268, 411-415, 1995
A:Title: FAP-1: a protein tyrosine phosphatase that associates with Fas.
A:Reference number: I59595; MUID:95232528
A:Accession: I59595
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1279-1888 <RES>
A:Cross-references: GB:L34563; NID:9806291; PIDN:AAC41755.1; PID:9806292
R:Maekawa, K.; Imagawa, N.; Negamatsu, M.; Harada, S.
FEBS Lett. 337, 200-206, 1994
A:Title: Molecular cloning of a novel protein-tyrosine phosphatase containing a membr
A:Reference number: I53483; MUID:94116679
A:Accession: I53483
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-61, 'GS', 64-839, 'D', 841-1210, 'T', 1212-1383, 1389-2299, 'QM', 2302-2490 <RE2
A:Cross-references: GB:D21209; NID:9452189; PIDN:BA04750.1; PID:9452190
C:Genetics:
A:Gene: GDB:PTPN13
A:Cross-references: GDB:306348; OMIM:600267
A:Map position: 4q21.3-4q21.3
C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 13; GLGF domain homolog
C:Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; tyr
F:574-868/Domain: protein 4.1 membrane-binding domain homology <B41>
F:1099-1175/Domain: GLGF domain homology <GLG1>
F:1373-1454/Domain: GLGF domain homology <GLG2>
F:1511-1590/Domain: GLGF domain homology <GLG3>
F:1799-1870/Domain: GLGF domain homology <GLG4>
F:1893-1967/Domain: GLGF domain homology <GLG5>
F:2242-2461/Domain: protein-tyrosine-phosphatase homology <PTP1>
F:2413/Active site: Cys (phosphocysteine intermediate) #status predicted
F:2419/Binding site: substrate phosphate (Arg) #status predicted

Query Match 7.18: Score 494.5; DB 1; Length 2490;
Best Local Similarity 22.66; Pred. No. 6.8e-19;
Matches 243; Conservative 140; Mismatches 371; Indels 319; Gaps 38;

```

Oy 530 EPPSLSIGSYGGRGMSRLNCEVMRGIFIKYHLEDSPGKNGTKLPGRIVAPQSGE 589
Db 990 EPPQTYAELVGRKSHOMNSDAESLAGV---TKLNSKSVASLNRSRPERKHESSSST 1046
Oy 590 SEPEKAPLCVPPPPSAFAEMGSDHTQSSASKISQDVNDEDEFGYSWKNIRERYGTLTG 649
Db 1047 EDPQAVYLVGKT-----MHSSGNSSSQVPLKENDV-----LHKRWSTVSS 1086
Oy 650 ---ELMHIELEK-GHSGLGSLAGNKRDR--MSVFIYGLIDPNCAGKDRLOIADLEL 703
Db 1087 PEREITLVNKKAKAYGIGFOITIGCEKMGRLDGIFFISSVAPGAPADLDCLCPGDRILIS 1146
Oy 704 INQILYGRSHONASSIIKCAPSKYIIIFIRNKDAVNOMVACGNANPELPSENENLOK 763
Db 1147 VNSVLEGVSHAAIEITLQNNAPEDVTLVISQPEKIKSVPTPVH---LTNEMKNYMKK 1202

```


Query Match	5.8%	Score 407.5	DB 2	Length 852
Best Local Similarity	28.4%	Pred No. 8.2e-15		
Matches 155	Conservative 68	Mismatches 212	Indels 111	Gaps 22

submitted to the EMBL Data Library, October 1995
A:Description: The sequence of C. elegans cosmid C25F6.
A:Reference number: 218377
A:Accession: J15617
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1131 <EEN>
A:Cross-references: EMBL:U039742; NID:g1049455; PID:g1049459; PIDN:AAA80434.1; CESP:C25F6
A:Gene: CESP:C25F6.2
A:Introns: 105/1; 210/3; 283/3; 316/1; 346/2; 463/3; 566/1; 722/2; 897/1; 991/2; 1032/2

Query Match 5.7%; Score 399.5; DB 2; Length 1131;
Best Local Similarity 24.4%; Pred. No. 3,7e-14;
Matches 185; Conservative 102; Mismatches 322; Indels 149; Gaps 27;

QY 494 DIPELPEREEGEGBS---ELONTAVSNMOPRRVELMREPSLSLGIYVGKMG SRL 549
Db 171 EVRRVRIETLEG-GHSHVSNRPAAATSTSNMVLSSSTPLISDLRDRGFSYLMGGGLNG L 229
QY 550 SNGEWMGIFLKHVLEDSPACKNGTLKPGDRIVEAPSOSESEPEKAPLCSPVPPPSAFA 609
Db 230 GNG-----LGNGLSSPNSSST---HYLHROROTSHDGWRRETTTIVTPPGLE 278
QY 610 EMGSDHTQSSASKISODVDKEDFEFGYSWKNIERYGYTLGELHMIIELEKSHGSLSLAG 669
Db 279 RRVVEHT-----GVIDHGRKW-----ELENIYVLEKHTGTGLGFSITYG 315
QY 670 NKDR-----SRMSVITVIGDPRNGAGKGRLOIDELLEINQILKXSHQMASSITICAP 725
Db 316 GMDPTEDEGDTSIYVTNIEGGAALADGRMKRNDIITAVNNTGCENKHEVAVALASSG 375
QY 726 SKVYIIFIRNKD-----AVNQAVACP-----GNAVEPLPSNSELNKETEPEP 767
Db 376 NVVSLSLKRRDEAFELPIGNGFGSTSYLRSQVPSVAGMLQAHHSAPRHPPEPP 435
QY 768 TVTTSDAVDSLRSK---NQVHLEPRKDOGGGLAIS-----EDDTLSGVIITSLT 815
Db 436 VHHGSLSQLSGOYRSTRPTNSYIDLKVGARGLFSIAGCGGNHNVKGDY-DIYVKII 493
QY 816 EHGVAATDGRKLVDOQLAIVDEIYGVPIEKFISLTKTAMVKTITHA-----ENP 868
Db 494 EEGAELDGRIRVGDKILVEDHSLINTTHNNAVNLKNGNRVRLLIQGTGALFNDSA 553
QY 869 DSGAVPSNAGAASEKKNSOSLWVPOS---GSPEPDSIRNTRSRSTPAIFASDPATCP 924
Db 554 SQQMPTPTPIRPSVODYNRSQMSQSHLSYGGP-----LNTSYSS-----CAPAIIP 602
QY 925 IIPCEITTEIISKRTGIGIYVGSSTLIGAFIIEHYEAGAACKRGRLMAGQIILEVN 984
Db 603 LEP---RPVLIKQGNLGNITVGGEDN--EPIITISYVLEGGVADLSGNKTKGDVLEVN 657
QY 985 GIDLRKATHDEAIVNLKOTPORVRLT-YRDEAPYKEEVCDTLIEIQRKPKGGLSTI 1043
Db 658 GVVLRNATHKEAALNAGNPVYLTLYQYRPQ---EQYFES---KIEKLNDVIAQSR 710
QY 1044 VGRKNDTGVVSDIVKGIADPDGR-----LIQGOILLVNGED-----VRN 1085
Db 711 MGLTSRSESEYVRLFD---YDPSHENSVAHRSMGFNYGDLIHINSSDDEMTARKVHE 767
QY 1086 ASQEAVALKCKSLGVTTLVEGRKAKGPFHS-----ERRPSQTSOVSGSL 1131
Db 768 NGERTAGAVIPSKKRVKRRRLRKQVNFNSGOSLSGRNSTTGLENKRSRSQLSFS-- 825
QY 1132 SSFTFPLSGSSTSESSLESSKKNALASEIOGLRPTVEMK 1169
Db 826 RKPFVFKSTDLRLNDLNESS--NVAEPPVMSYQAVEGO 861

RESULT 14
114152
synaptic scaffolding protein S-SCAM - rat

Query Match 5.7%; Score 399.5; DB 2; Length 1131;
 Best Local Similarity 24.4%; Pred. No. 37e-14;
 Matches 185; Conservative 102; Mismatches 322; Indels 149; Gaps 27;

QY	494	DIPELPEREAGEEES	---	ELQNTAVSNMNP	PRVELMREPS	SLGISYVGKMG	SRL	549
		:: :: :: :: ::		:: :: :: :: ::				
DB	171	EVRRVIERLEEG	-GPHSYNSRPAAT	STSNMYLS	STTPI	DLDRGCGF	SYLNGG	1229
QY	550	SNGEVMGICFI	KHNL	EDSPACKNG	TILKPGD	IVAPSP	SESEPEKAP	LCSPPPP
		:: :: :: :: ::		:: :: :: :: ::		:: :: :: :: ::		
DB	230	GNG	---	IGNLIS	SPYNS	ST---	HYLHRQR	QTS
QY	670	NKDR	---	SRMSV	ITGIDP	NGAAGK	DRLQ	IAD
		:: :: :: :: ::		:: :: :: :: ::		:: :: :: :: ::		
DB	316	GMDP	PTEDG	DTIS	IVTNI	IEG	GALADGR	MRKND
QY	726	SKVY	IIIRND	-----	AVNQA	AVCP	---	GNA
		:: :: :: :: ::		:: :: :: :: ::		:: :: :: :: ::		
DB	376	NVSL	SLKRRDE	AFLE	PGNFG	SGST	LSNG	CVPS
QY	768	TVTTS	DAVD	LSFK	---	NVQHL	EP	RD
		:: :: :: :: ::		:: :: :: :: ::		:: :: :: :: ::		
DB	436	VHHS	LSQL	SVGQ	RS	TRP	MTSV	IDL
QY	816	EHG	VAADGR	KAVD	QILAND	DEL	IVG	PI
		:: :: :: :: ::		:: :: :: :: ::		:: :: :: :: ::		
DB	494	EEGA	AEELDGR	LAVG	DKILE	VDHSL	INT	TH
QY	869	DSQ	AVPSA	GAAS	GEK	KNSS	OL	WPOS
		:: :: :: :: ::		:: :: :: :: ::		:: :: :: :: ::		
DB	554	SQ	MP	PTPI	LR	BS	VOD	Y
QY	925	IIP	CE	TTI	EL	IS	KRT	GI
		:: :: :: :: ::		:: :: :: :: ::		:: :: :: :: ::		
DB	603	LEP	---	RP	YLV	KQNG	IG	FN
QY	985	GID	LK	AT	H	DE	AL	IV
		:: :: :: :: ::		:: :: :: :: ::		:: :: :: :: ::		
DB	658	G	VY	LR	NR	AT	HE	AE
QY	1044	V	K	R	N	D	I	G
		:: :: :: :: ::		:: :: :: :: ::		:: :: :: :: ::		
DB	711	M	G	T	S	R	K	S
QY	1086	A	S	O	E	A	V	A
		:: :: :: :: ::		:: :: :: :: ::		:: :: :: :: ::		
DB	768	N	G	E	T	A	G	V
QY	1132	S	S	F	P	L	S	G
		:: :: :: :: ::		:: :: :: :: ::		:: :: :: :: ::		
DB	826	R	K	P	F	V	K	S
		:: :: :: :: ::		:: :: :: :: ::		:: :: :: :: ::		

RESULT 14
 T14152
 dynamic scaffolding protein S-SCAM - rat

C.Species: Rattus norvegicus (Norway rat)
 C.Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Aug-2000
 C.Accession: T14152
 R.Hitao, K.; Hata, Y.; Ide, N.; Takeuchi, M.; Irie, M.; Yao, I.; Deguchi, M.; Toyoda, A.
 J. Biol. Chem. 273, 21105-21110, 1998
 A.Title: A novel multiple PDZ domain-containing molecule interacting with N-methyl-D-aspartate receptor
 A.Reference number: Z17889; MUID:98361985
 A.Accession: T14152
 A.Status: Preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: mRNA
 A.Residues: 1-1277 <HIT>
 A.Cross-references: EMBL:AF034863; NID:g3411052; PID:g3411053; PIDN:AAC31124.1
 C.Genetics:
 A.Gene: S-SCAM
 A.Function:
 A.Description: may assemble receptors and cell adhesion proteins at synaptic junctions
 C.Superfamily: WW repeat homology
 F:302-339/Domain: WW repeat homology <WM1>
 F:348-385/Domain: WW repeat homology <WM2>

Query Match 5.7%; Score 398.5; DB 2; Length 1277;
 Best Local Similarity 19.9%; Pred. No. 5.1e-14;
 Matches 221; Conservative 146; Mismatches 333; Indels 413; Gaps 46;

QY 30 PHVDGEFISSEDEDPVLAAMDAGOSTEEVOAP-----LAMWAGIHIIELEK 78
 DB 393 PTELEA-----TKPLQAPGFRERPLETRDASQLKGFLLSTTLKK 431

QY 79 GSKGLGFSILDYODPIDPASTVILIRSLVPGIAEKGRLLPGRILMFVNDLENLSLE 138
 DB 432 SMWGFETIIGDEP---DEFLQVKSVPIDGPAAQDKMERTGDIYVINEVCVGHHA 487

QY 139 EAVEALKAPSG-TYRIGYAK--PLPLSPEGVSAKEDSFYIPHSCEAGIADKPLER 195
 DB 488 DYVKLEQSVPIQSVNLVYCRGPPLFDEPDSPANS-----WVPP-----LAIMERP--- 533

QY 196 ADLALVGTADADLVDESPESPSPENDSIYSQASILSLHSSCGDGLNYSGLSSPP 255
 DB 534 -----PRVMNGRINNYETLEYISRTSOSVDP-----ITDRPP 566

QY 256 KDVIENSCDPVLDLHMSLELYTONLIERODENTPVSIDSMGPASGFTINDYTPANAIDQ 315
 DB 567 H-----SLH-----SM-PADQDLGTTPP----- 585

QY 316 QYECENTIVTWSHLPSEVYSSALPVLPSDAGKSEHLLRQSSSLACAEVMAQNVSK 375
 DB 586 -----VHDDNVSYA-----SSGATQAEILM----- 604

QY 376 ESFERTINAKNSLSGMVYVANKDGLMIVRSIIHGCAISNDGRILAIDCIIISINEST 435
 DB 605 -----TLTIVKAKGFGFTIADSP--TGQRVQOILD--IQGCPGLCEBDLVEINQOVN 654

QY 436 ISVTNAQARMLRRHSLIPDIKITVYPAHLEEFKISLGQSGRVMALDIFSSYGRDI 495
 DB 655 QNLSTHE-----VVDIL-----KVC 669

QY 496 PELPEREGEGESEL--ONTAVSNMNPOR-RVELMR--EPEKSLGISVGGRMGSR 548
 DB 670 P-----VGSETSLIIRGFFSPMKPKPMVDHMEQNGSPQSLSPAPV----- 713

QY 549 LNGEVMRGIIFIKHVLDEDPACKNGTLKPGDRIVEAPSSSES-EPEKALCISVPPRPPSA 607
 DB 714 -----PQSLPPPALHRSF-----PDSTEARDPK-----PDPE 744

QY 608 PAEMGSDHTOSSASKISQDVDEKDEFGYSKMNIRERYGTLTELMHIELEKSHGLSL 667
 DB 745 LKESKATIESRQVPPRISFRDSSGPRYKEL-----DVHLRRME--SGGFRFI 792

QY 668 AGNKRDSRNSVPIYGDIPGAAGKGRQLADLELEINQOILYGRSHQANASIIKCAPSK 727
 DB 793 LGG-DEPGQPLITGAVIANGSADRDGRILHPGDELVYVDGIPVAGKTHRYVIDLMHHAARN 851

QY 728 VKI-IFIRNKDANVMAYCPGNAVEPLPSNSNLONKETEPIYTTSDAAVDSLSPKNQV 786
 DB 852 GGVNLTIVRRKK-----VLGG-----EPCPENGSRPSGVSTHSSPRSDYA---TVANSNH 898

QY 787 LELPKDQG-----LGIAISEDTLGGYIKSLT--EHGVAAT----- 822
 DB 899 -AAPSNMNSPREGFASHSLQSDYIIHKKENGSGFVYISSLNRESGATTVIPKIGRI 957

QY 823 -DG-----RLKVDQILAVDEIVGVPIEKETSLKTAKMTVKLT---HAENPDS 870
 DB 958 IDGSPADRCALKVDDRIILVNGOSIIMPHADYKLIKDAOLSTLRIIPQEEINPPTS 1017

QY 871 QAVPSAGASGEKKN-----SSQSLMVPQSSPEDESIR---N 906
 DB 1018 -----APSEKQSPMAQHSPLAQHSPLAQSPPATPNSPVAQAPAPQLQLQGHEN 1069

QY 907 TSSRSTPA-----IFASDPATCPPIEGCE----- 930
 DB 1070 SYRSEVKANQDVKPDIRQPFPTDYRQPLDYRQPPGQVSGSPPLDYRQHSPTQRYPLS 1129

QY 931 -----TTIEISKRTGLGLSVGSDTLGAFIIEHYEERGACKGRILNAGDI 980
 DB 1130 DYROPQDEDFYFVDMKGAKGFGFSIRGREGYKMDLYVL-RLAEDGPALIRGRMVGDI 1188

QY 981 LEVNGIDLRKATHDPAINVLROTPOVRRLTLR 1013
 DB 1189 IEINGESTDMTHARALEIKSGRRVRLLLKR 1221

RESULT 15
 G01974
 channel associated protein of synapse - human
 C.Species: Homo sapiens (man)
 C.Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 21-Jan-2000
 C.Accession: G01974
 R.Kim, E.; Cho, K.
 submitted to the EMBL Data Library, July 1995
 A.Reference number: G08966
 A.Accession: G01974
 A.Status: preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: mRNA
 A.Residues: 1-870 <KIM>
 A.Cross-references: EMBL:U32376; NID:g1463025; PID:g1036790
 C.Genetics:
 A.Gene: chapsyn-110
 C.Superfamily: discs-large tumor suppressor; GLGF domain homology; guanylate kinase h
 F:198-276/Domain: GLGF domain homology <GLG2>
 F:543-601/Domain: SH3 homology <SH3>
 F:681-858/Domain: guanylate kinase homology <GKI>

Query Match 5.7%; Score 397.5; DB 2; Length 870;
 Best Local Similarity 28.1%; Pred. No. 3.1e-14;
 Matches 154; Conservative 69; Mismatches 205; Indels 121; Gaps 23;

QY 868 PQSQAIVPSAAGASC-----EKKNS-----QSLMVPQSSPEDESIRTSRS 911
 DB 24 PDDHSLPRLTHEVRGPELVHSEKMLQLENVHGYVLOSHISPLKASAP-ITVMDTL 82

QY 912 TPAIFASDPATCPPIEGCEFTIE-----ISKRTGLGLSVGSDPT-LG---AFIIEHY 962
 DB 83 -----TIPYVNGTIEIEFEETILERKNSGLGFSIAGTDNPHIGDDPGIFITRI 132

QY 963 YEEGAACKDRLMAGDQILLEVNGIDLRKATHDPAINVLROTPOVRRLTLRYDEAPYKEE 1022
 DB 133 IFGGAIAEEDGRILRVNDICIRVNEVDVSEVSHSKAVALKEASIRALTY-RRRRPILTEV 191

QY 1023 VCDTLTIELQKPGKGLGISTYK-----RNDTGVPVSDIYKGGIADPDGRLIGDQIL 1076
 DB 192 V-----EIKLFKRP-KGLGFSIAGVGNQHIIPDNSIYVKRIIDGGAAGQDGRILQVGDRL 246

QY 1077 LVNGEDVRNASQEAVALKCSIGVTLLEVGRKA-----GP---FHSERRPQSOTSOY 1126
 DB 1077 LVNGEDVRNASQEAVALKCSIGVTLLEVGRKA-----GP---FHSERRPQSOTSOY 1126

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2001, 14:39:47 ; Search time 21.71 Seconds

(without alignments)
2166.411 Million cell updates/sec

Title: US-09-502-698-1

Perfect score: 6990

Sequence: 1 MWCRRTPPTTQSELDLD.....EEAVAIKKRTKGTITLVLS 1373

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	509	7.3	2485	1 PTND_HUMAN	Q12923 homo sapien
2	405.5	5.8	852	1 DTG2_RAT	Q63622 rattus norv
3	399	5.7	817	1 DTG3_HUMAN	Q92796 homo sapien
4	397.5	5.7	870	1 DTG2_HUMAN	Q15700 homo sapien
5	393	5.6	849	1 DTG3_MOUSE	P70175 mus musculu
6	391	5.6	849	1 DTG3_RAT	Q62936 rattus norv
7	384	5.5	767	1 DTG4_HUMAN	P78352 homo sapien
8	383	5.5	724	1 DTG4_RAT	P31016 rattus norv
9	382	5.5	724	1 DTG4_MOUSE	Q62108 mus musculu
10	379	5.4	960	1 DTG1_DROME	P31007 drosophila
11	378.5	5.4	911	1 DTG1_RAT	Q62696 rattus norv
12	377	5.4	904	1 DTG1_HUMAN	Q12959 homo sapien
13	283	4.0	631	1 I116_HUMAN	Q14005 homo sapien
14	243	3.5	1736	1 ZOI_HUMAN	Q07157 homo sapien
15	242	3.5	933	1 ZOI_HUMAN	Q95049 homo sapien
16	240	3.4	898	1 ZOI_CANFA	Q62683 canis famli
17	237.5	3.4	1745	1 ZOI_MOUSE	P33447 mus musculu
18	235.5	3.4	905	1 ZOI_MOUSE	Q94971 mus musculu
19	211	3.0	1816	1 AFE_HUMAN	P55196 homo sapien
20	205.5	2.9	839	1 AFE_RAT	Q35430 rattus norv
21	197.5	2.8	3562	1 PCG3_CHICK	Q90953 gallus gall
22	185.5	2.7	2109	1 PCG3_CHICK	P07898 gallus gall
23	184.5	2.6	1972	1 P531_HUMAN	Q12888 homo sapien
24	182.5	2.6	982	1 ABPA_CAEEL	Q17583 caenorhabdi
25	182	2.6	206	1 OM25_RAT	Q9wvj4 rattus norv
26	178	2.5	2124	1 PGCA_RAT	P07897 rattus norv
27	177.5	2.5	571	1 APB3_MOUSE	Q08888 mus musculu
28	177	2.5	837	1 APB3_HUMAN	Q02410 homo sapien
29	176	2.5	145	1 OM25_HUMAN	P57105 homo sapien
30	173.5	2.5	569	1 APB3_RAT	Q70248 rattus norv
31	173.5	2.5	750	1 APB2_RAT	Q33431 rattus norv
32	173	2.5	2468	1 MAPB_HUMAN	P46821 homo sapien
33	173	2.5	3358	1 PCGV_MOUSE	Q62059 mus musculu

ALIGNMENTS

RESULT	ID	PTND_HUMAN	STANDARD	PRT: 2485 AA	Q01282 mus musculu
AC	Q12923	Q15263; Q16826; Q15264; Q15265; Q15159;	01-OCT-2000 (Rel. 40, Created)	736 1	Q60838 mus musculu
DT	01-OCT-2000 (Rel. 40, Last sequence update)			2842 1	P70478 rattus norv
DT	01-OCT-2000 (Rel. 40, Last annotation update)			2845 1	Q61315 mus musculu
DE	PROTEIN-TYROSINE PHOSPHATASE, NON-RECEPTOR TYPE 13 (EC 3.1.3.48)			1845 1	Q9u136 homo sapien
DE	(PROTEIN-TYROSINE PHOSPHATASE 1E) (PTP-ET) (PTP-BAS) (PROTEIN-TYROSINE PHOSPHATASE PTP1) (FAS-ASSOCIATED PROTEIN-TYROSINE PHOSPHATASE 1) (FAP-1).			1608 1	P15320 seretelia ma
DE	PTPN13 OR PTP1E OR PTP1 OR PNP1.			2314 1	P23471 homo sapien
GN	Homo sapiens (Human).			3396 1	P25054 homo sapien
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			2843 1	Q03133 saccharopol
OC	NCBI-Taxid:9606;			3172 1	P16466 proteus mir
OX	[1]			1577 1	Q02952 homo sapien
RP	SEQUENCE FROM N.A.			1781 1	
RC	TISSUE-Breast carcinoma;				
RX	MEDLINE-94350988; PubMed-8071359;				
RA	Bayville D., Ahmed S., Stocco R., Shen S.-H.;				
RT	"A novel protein-tyrosine phosphatase with homology to both the cytoskeletal proteins of the band 4.1 family and junction-associated guanylate kinases.";				
RL	J. Biol. Chem. 269:22320-22327(1994).				
[2]	SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.				
RC	TISSUE-Leukemia;				
RX	MEDLINE-94116679; PubMed-8287977;				
RA	Maekawa K., Imagawa N., Nagamatsu M., Harada S.;				
RT	"Molecular cloning of a novel protein-tyrosine phosphatase containing a membrane-binding domain and GlGF repeats.";				
RL	FEBS Lett. 337:200-206(1994).				
[3]	SEQUENCE FROM N.A.				
RC	TISSUE-Fibroblast;				
RX	MEDLINE-95014139; PubMed-7929060;				
RA	Saris J., Claesson-Welsh L., Heldin C.-H., Genez L.J.;				
RT	"Cloning and characterization of PTP1, a protein tyrosine phosphatase with similarities to cytoskeletal-associated proteins.";				
RL	J. Biol. Chem. 269:24082-24089(1994).				
[4]	SEQUENCE OF 1216-2490 FROM N.A.				
RC	TISSUE-Pancreas;				
RA	Submitted (JUN-1994) to the EMBL/Genbank/DBJ databases.				
RL	Wang H. Y.;				
[5]	STRUCTURE BY NMR OF 1361-1456.				
RP	MEDLINE-20170882; PubMed-10704206;				
RX	Kozlov G., Gehring K., Ekkel I.;				
RA	"Solution structure of the PDZ domain from human phosphatase hPTP1E and its interactions with C-terminal peptides from the Fas receptor.";				
RL	Biochemistry 39:2572-2580(2000).				
CC	-1- FUNCTION: BINDS TO A NEGATIVE REGULATORY DOMAIN IN FAS THAT INHIBITS FAS-INDUCED APOPTOSIS.				

OY 1213 GDRIVTIGTSTEGMTHQAVNLKKNASGSIEMOVAGGVSVTHGHOE-PASSLSFT 1271
 DB 1831 GRLKLVNDVTYNTMTHTDAVNLRAASTVVL-----VIGRVLELPRIPML--- 1877
 OY 1272 GLTSTSTPDDGLPPQCKSTLERGPDGIVSGVSPHGLDLYVTVFAKGAASD 1331
 DB 1878 -----PHLLPDTITLCKKEELGFSLCGSHDLV--QVYVLSIDINPRVAALE 1922
 OY 1332 GLKRGDQIIAVNGSLGCVTHEEA-----VAIKRTGVTYLM 1370
 DB 1923 GMLQLDLVHYVNGVSTQGMTELEVNRALDMSLPISLVLKATRNLDLPV 1970
 RESULT 2
 DLG2_RAT 2
 ID DLG2_RAT STANDARD: PRT: 852 AA.
 AC Q63622: Q62939; P70548;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE CHANNEL ASSOCIATED PROTEIN OF SYNAPSE-110 (CHASYN-110) (SYNAPTIC DENSITY PROTEIN PSD-93) (DISCS, LARGE HOMOLOG 2).
 GN DLG2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96310881; PubMed=8755482;
 RA Kim E., Cho K.O., Rothschild A., Sheng M.;
 RT "Heteromultimerization and NMDA receptor-clustering activity of
 RL Chapsyn-110, a member of the PSD-95 family of proteins";
 RN Neuron 17:103-113(1996).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96193770; PubMed=8625413;
 RA Breman J.E., Chao D.S., McGee A.W., Craven S.E.,
 RA Santillano D.R., Wu Z., Huang F., Xia H., Peters M.F.,
 RA Froehner S.C., Bredt D.S.;
 RT "Interaction of nitric oxide synthase with the postsynaptic density
 RL protein PSD-95 and alpha1-syntrophin mediated by PDZ domains.";
 RN Cell 84:757-767(1996).
 [3]
 RP SEQUENCE FROM N.A.
 RA Irie M., Hata Y., Takai Y.;
 RL Submitted (Sep-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR
 CC SUBUNIT AS WELL AS POTASSIUM CHANNELS.
 CC -1- SIMILARITY: CONTAINS 3 PDZ/DRH DOMAINS.
 CC -1- SIMILARITY: CONTRAINS A GUANYLATE KINASE-LIKE DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U49049; AAB53243.1; -
 DR EMBL: U50717; AAC52643.1; -
 DR EMBL: U53368; AAB48562.1; -
 DR HSP: Q12959; IPDR.
 DR InterPro: IPR000619; -
 DR InterPro: IPR001452; -
 DR InterPro: IPR001478; -
 DR Pfam: PF00625; Guanylate_kin; 1.
 DR Pfam: PF00595; PDZ; 3.
 DR Pfam: PF00018; SH3; 1.
 DR PROSITE: PS00856; GUANYLATE_KINASE_1; 1.

DR PROSITE: PS0052; GUANYLATE_KINASE_2; 1.
 DR PROSITE: PS0106; PDZ; 3.
 DR PROSITE: PS0002; SH3; 1.
 KM SH3 domain; Repeat.
 FT DOMAIN 98 PDZ 1.
 FT DOMAIN 193 PDZ 2.
 FT DOMAIN 421 PDZ 3.
 FT DOMAIN 536 SH3.
 FT DOMAIN 662 SH3.
 FT DOMAIN 852 GUANYLATE KINASE.
 FT CONFLICT 181 VR -> IL (IN REF. 2).
 FT CONFLICT 228 I -> M (IN REF. 2).
 FT CONFLICT 326 R -> K (IN REF. 2).
 FT CONFLICT 339 D -> E (IN REF. 3).
 FT CONFLICT 450 MISSING (IN REF. 2).
 FT CONFLICT 454 GD -> RK (IN REF. 2).
 FT CONFLICT 464 D -> H (IN REF. 2).
 FT CONFLICT 474 R -> P (IN REF. 2).
 FT CONFLICT 476 A -> D (IN REF. 2).
 FT CONFLICT 478 AAA -> LP (IN REF. 2).
 FT CONFLICT 484 A -> S (IN REF. 2).
 FT CONFLICT 506 H -> N (IN REF. 2).
 FT CONFLICT 569 L -> Q (IN REF. 2).
 FT CONFLICT 586 MISSING (IN REF. 2).
 FT CONFLICT 626 K -> A (IN REF. 3).
 FT CONFLICT 639 F -> L (IN REF. 1).
 FT CONFLICT 726 N -> Y (IN REF. 2).
 FT CONFLICT 733 E -> V (IN REF. 1).
 FT CONFLICT 749 L -> H (IN REF. 2).
 FT CONFLICT 756 KR -> NG (IN REF. 2).
 FT CONFLICT 791 T -> M (IN REF. 2).
 FT CONFLICT 794
 SO SEQUENCE 852 AA; 94934 MM; FBD414AB9CF5B09 CnC64;
 Query Match 5.8%; Score 405.5; DB 1; Length 852;
 Best Local Similarity 28.4%; Pred. No. 4.6e-14;
 Matches 155; Conservative 67; Mismatches 213; Indels 111; Gaps 22;
 OY 866 EMPDQAVPSAAGASG-----EKKNS-----QSLMVPQSGSPEDSIRNRS 909
 DB 22 DGPDHSLRLTHEYRGPELVHSEKNSQIENVHGYLSHSPKASPAV-IIVNTDT 80
 OY 910 SSTPAIFASDPATCPPIPCETTIE-----ISKRTGLGLSVGSDT-LIG---AFIIN 960
 DB 81 LD-----TIYVNTET 130
 OY 961 EYEEGACKDGRILWAGDQILEVNGIDLRKATHEAINVLRKTPORVRLTYRDEAPYE 1020
 DB 131 KIIPGGAALDGRILRVNDCILRVNEVDVSEVSHKAVELAKGASIVRLYV-RRRRPIL 189
 OY 1021 EYVCDTLTLELKKKKGKGLSVYK-----RNDTGVFVSIVVGGIADPGRLIOGDO 1074
 DB 190 TVV-----ETLKRGP-KGLGFSIAGVGNQHIIPGDSITVYKIIDGGAQKGRLOVGR 244
 OY 1075 ILLVNGEDVRNASEVAALLKCSLGTVLLEVGRKA-----GP---FHSERRPSQTS 1124
 DB 245 LLMVNYSLEEVTHEBAVAIILKNTSDVYLVKGRPTIYMTDPPGPDITHSYSPMEWH 304
 OY 1125 QVSEGLSFTPLPGSSISSELESSKKNALASELQIGRTYEMAKGPPDSIGISAGV 1184
 DB 305 LLS-GNGNGLEYKTSLLPISPRYSPIPKMLV-EDDYTRPE---PVTSTVKNLCCKP 358
 OY 1185 GSPLDGVPIFIIMMHPITGVAAGTOKLRVGDRIIVTIGTSTEGMTHQAVNLKKNASGSI 1244
 DB 359 ASPRKHSVP-----ECKDSF-----LSTYPPNH 383
 OY 1245 MOVVAGGVSVYTHGHOEPASSLSFTGLTSTSTPDDGLPPQCKSTLERGPDGLGFSI 1304
 DB 384 LGLLPDSMTSHSQSTATRGPSVTLQRAISLE-----GEP--RKVYLHKSGTGIFNI 435
 OY 1305 VCGYSGPHDLPITYKTVAKGAASEDGLKRGDQIIAVNGSLGCVTHEEVALIKRK 1364
 DB 436 VGG---EDGE-GIFVSFIAGPADLSELQRGDQILSVNGSIDLKGASHEDQAAALKGAG 491

DB	QTVTII	497	oy	1365	GVTTLM	1370
:						
RESULT	3					
DLG3_HUMAN	STANDARD;	PRT;	817	AA.		
AC	093796;	OSU18;				
DT	01-NOV-1997	(Rel. 35, Created)				
DT	01-NOV-1997	(Rel. 35, Last sequence update)				
DE	01-OCT-2000	(Rel. 40, Last annotation update)				
DE	PRESYNAPTIC PROTEIN SAPI02	(SYNASE-ASSOCIATED PROTEIN 102)				
GN	(NEUROENDOCRINE-DLG)	(NE-DLG) (DISCS, LARGE HOMOLOG 3).				
OS	DLG3 OR KIAA1232.					
OC	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
OX	NCBI_TaxID=9606;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE=Fetal brain;					
RX	MEDLINE=97332623; PubMed=9188857;					
RA	Makino K., Kuwahara H., Masuko N., Nishiyama Y., Morisaki T.,					
RA	Sasaki J., Nakao M., Kuwano A., Nakata M., Ushio Y., Saya H.;					
RT	"Cloning and characterization of NE-dlg: a novel human homolog of the					
RT	Drosophila discs large (dlg) tumor suppressor protein interacts with					
RL	the APC protein.";					
RL	Oncogene 14:2425-2433(1997).					
RN	[2]					
RP	SEQUENCE OF 330-817 FROM N.A.					
RC	TISSUE=Brain;					
RX	MEDLINE=20033619; PubMed=10574462;					
RA	Nagase T., Ishikawa K.-I., Kikuno R., Hiroseawa M., Nomura N.,					
RA	Ohara O.;					
RT	"Prediction of the coding sequences of unidentified human genes. XV.					
RT	The complete sequences of 100 new cDNA clones from brain which code					
RT	for large proteins in vitro.";					
RL	DNA Res. 6:337-345(1999).					
CC	-1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF THE NMDA RECEPTOR					
CC	SUBUNIT NR2B (BY SIMILARITY).					
CC	-1- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.					
CC	-1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.					
CC	-1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.					
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/					
CC	or send an email to license@isb-sib.ch).					
CC	-----					
DR	EMBL; U49089; AAB61453.1; -.					
DR	EMBL; AB033058; BAA86546.1; -.					
DR	HSSP; Q12959; 1PDR.					
DR	MIW; 300189; -.					
DR	InterPro; IPR000619; -.					
DR	InterPro; IPR001452; -.					
DR	InterPro; IPR001478; -.					
DR	Pfam; PF00625; Guanylate_kin; 1.					
DR	Pfam; PF00595; PDZ; 3.					
DR	Pfam; PF00018; SH3; 1.					
DR	PROSITE; PS00856; GUANYLATE_KINASE_1; 1.					
DR	PROSITE; PS00552; GUANYLATE_KINASE_2; 1.					
DR	PROSITE; PS0106; PDZ; 3.					
DR	PROSITE; PS50002; SH3; 1.					
KM	SH3 domain; Repeat.					
FT	DOMAIN 130 217 PDZ 1.					
FT	DOMAIN 226 311 PDZ 2.					
FT	DOMAIN 379 465 PDZ 3.					
FT	DOMAIN 503 568 SH3.					

FT	DOMAIN	628	803	GRANULATE KINASE
FT	CONFLICT	330	381	FRALADNTE SHMSLGIYGAIVESKYSYPAPQVPPTRKSP
FT	CONFLICT	592	606	PRHMLAEEDFT->AAREKERAMERAKESGSLANGLG
FT	CONFLICT	592	606	ASSASARRASQWAMPRLSLRGDA (IN REF. 2).
FT	CONFLICT	592	606	DEPGSLDDYDGAKNL->SITKTRKSPFLSRPFYKS
FT	CONFLICT	592	606	ENNAOESSIOEGVTSNTSDSESS (IN REF. 2).
FT	SEQUENCE	817 AA;	90344 MW;	3D7512BCA413FC4E CRC64;
SO	SEQUENCE	817 AA;	90344 MW;	3D7512BCA413FC4E CRC64;
Query Match	Best Local Similarity	5.78;	Score 399;	DB 1; Length 817;
Matches 132;	Conservative	77;	Mismatches 168;	Indels 154; Gaps
OY	870	SOAPVSAAGAASGEKKNSSOSLWPOSGSPESPESINTSRSSPTAFADPATPCIPGC	929	
DB	56	SOTLPSSQAGATPPTPR---TKAKLIPTRGRDVPPLPKPVGCKSTPKLNGSGPFW---	2EC 109	
OY	930	ETT-----LEISKRTGLGLSIVGSDTL-----LGAFIIEHYEE	965	
DB	110	TCTNRDWEQVNGSDGFKFKEEYIELERNGSLGFSIAGIDNPVDDPGIRIT-KIIPG	168	
OY	966	GAACKDGRILWAGDQILLEVNIDLRKATHDAIVLWQTPQRRLTILXREAPYKEEVED	1025	
DB	169	GAALMDGRGLVNDVCLVNEVEVEVHSAVLALEAGPVRLVLRKQPPPE-----	T 223	
OY	1026	TLTTELQKKPKPGKGLSIVSK-----RNDTGVFSDIYKGIADPDGRILQDQILLVN	1079	
DB	224	IMEVNLKGP-KGLGFSIAGIGNQNHIPGNSIYITKIIEGGAQKDSGLQGRDLAVLN	282	
OY	1080	GEDVRNMSQEAVALLCKSLGTYTLEVGRITAKGPFHSERRPSPOTSQVSEGLSSTFP	1139	
DB	283	NTNLDQVRHEEVAASLNTSDMYLKYA-KPQSLH-----LMDYARP	325	
OY	1140	GSTSELESSSKNALALASEIQLRIVEMKKGPDTSLGISIAGVGSPGLGDPVFIAMH	1199	
DB	326	YASTFTALADNH-----	337	
OY	1200	PTGVAAOTOKLRWGRDRIVTTCSTEGMHTQAVNLNKKASSIEMQVAVGDSVVTGH	1259	
DB	338	-----ISHNSSLGTL--GAIVESK-----SY	356	
OY	1260	HQEPASSLSFTGLTSTSIIPDOLGPPCKOSTILEGPGGLSIVGSGPHGDLPIYV	1319	
DB	357	PAPQVPPTRKSPTRPRMLAEEDF-TREPRKILLKSGSLGLENIVYG---EDGE-GIFV	411	
OY	1320	KTVFAKGAASEDEGRKRGDQILIAVNGOSLEGVTHEEVAVALLRTKGYVTLM	1370	
DB	412	SFLAGBPADLSGELRGDRITLSVNGVNLNATHQDMAALKRAGQSVYIV	462	
RESULT	4			
DLG2	HUMAN	STANDARD;	PRT;	870 AA.
ID	DLG2	HUMAN		015700;
AC	015700;			
DT	01-NOV-1997	(Rel. 35, Created)		
DT	01-NOV-1997	(Rel. 35, Last sequence update)		
DT	01-OCT-2000	(Rel. 40, Last annotation update)		
DE	CHANNEL ASSOCIATED PROTEIN OF SYNAPSE-110 (CHAPSIN-110) (DISCS, LARGE			
DE	HOMOLOG 2).			
GN	DLG2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96310881; PubMed=8755482;			
RA	Kim E., Cho K.-O., Rothschild A., Sheng M.;			
RT	"Heteromultimerization and NMDA receptor-clustering activity of			
RT	Chapsin-110, a member of the PSD-95 family of proteins."			
RL	Neuron 17:103-113(1996).			
CC	-1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR			
CC	SUBUNITS AS WELL AS POTASSIUM CHANNELS.			

Query Match	5.7%	Score 397.5;	DB 1,	Length 870;
Best Local Similarity	28.1%	Pred. No. 1.2e-13;		
Matches 154; Conservative	69;	Mismatches 205;	Indels 121;	Gaps 23;

```
D0      433 ENIYGC---EDGE-GIFPSFLLAGPADLSEELRGQDIISVNGIDLRGASHEDQAALAK 488
OY      136Z RTKGVTILM 1370
Db       |::|:
         489 GAGQTWITII 497

RESULT    5
DLG3_MOUSE ID DLG3_MOUSE STANDARD; PRT; 849 AA.
AC          P701N75;
DT     01-NOV-1997 (Rel. .35, Created)
DT     01-NOV-1997 (Rel. .35, Last sequence update)
DE     01-OCT-2000 (Rel. .40, Last annotation update)
PR     PRESYNAPTIC PROTEIN SAPI02 (SIMPASE-ASSOCIATED PROTEIN 102) (DISCS, DE LARGE HOMOLOG 3).
GN     DLG3 OR DLGH3.
OS     Mus musculus (Mouse).
OC     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RX     NCBI_TaxID=10090;
[1]
RN      SEQUENCE FROM N.A.
RP      STRAIN=C57BL/6; TISSUE=Brain;
RC      Kohmura N., Makino S., Yagi T.;
RL      Submitted (Aug-1996) to the EMBL/GenBank/DDBJ databases.
CC      -1 FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF THE NMDA RECEPTOR SUBUNIT NR2B.
CC      -1 SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
CC      -1 SIMILARITY: CONTAINS A SH3 DOMAIN.
CC      -1 SIMILARITY: BELONGS TO THE MAGOK FAMILY OF CELL JUNCTION PROTEINS.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to ilicense@isb-sib.ch).
-----
CC      EMBL: D87117; BAA13249.1; -.
DR      HSSP; I02859; IPDR.
DR      MGd; MGt:1888986; DigH3.
DR      InterPro; IPR000619; -.
DR      InterPro; IPR001452; -.
DR      InterPro; IPR001478; -.
DR      Pfam; PF006625; Guanylate_kin_1.
DR      Pfam; PF00595; PDZ_3.
DR      Pfam; PF00018; SH3_1.
DR      ProSite; PS00856; GUANYLYATE_KINASE_1; 1.
DR      ProSite; PS50052; GUANYLYATE_KINASE_2; 1.
DR      ProSite; PS50106; PDZ_3.
DR      ProSite; PS50002; SH3_1.
SH3 domain repeat.
KW      SH3 domain; Repeat.
FT      FT              149           235             PDZ 1.
FT      DOMAIN          244           330             PDZ 2.
FT      DOMAIN          404           484             PDZ 3.
FT      DOMAIN          519           589             SH3.
FT      DOMAIN          659           849             GUANYLYATE KINASE.
SQ      SEQUENCE   849 AA;  93482 MW;  EF3EF2D513538BE CRC64;

Query Match               5.6%; Score 393; DB 1; Length 849;
Best Local Similarity     24.7%; Pred. NO. 2.le-13;
Matches 139; Conservative 73; Mismatches 173; Indels 178; Gaps 18;
```



```

RESULT 7
ID      DLG4_HUMAN      STANDARD:      PRT:      767 AA.
AC      P78352; Q92941; Q9UKK8;
DT      01-NOV-1997 (Rel. 35; Created)
DT      30-MAY-2000 (Rel. 39; Last sequence update)
DT      01-OCT-2000 (Rel. 40; Last annotation update)
DE      PRESYNAPTIC DENSITY PROTEIN 95 (PSD-95) (DISCS, LARGE HOMOLOG 4).
GN      DLG4 OR PSD95.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_Taxid=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Mammary gland;
RC      MEDLINE=97432822; PubMed=9286702;
RA      Stachakis D.G., Hoover K.B., You Z., Bryant P.J.;
RT      "Human postsynaptic density-95 (PSD95): location of the gene (DLG4)
RT      and possible function in nonneural as well as in neural tissues.";
RL      Genomics 44:71-82(1997).
RN      [2]
RP      REVISIONS.
RC      TISSUE=Mammary gland;
RC      Stachakis D.G., Hoover K.H., You Z., Bryant P.J.;
RL      Submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RA      Stachakis D.G., Udar N., Sandgren O., Andreasson S., Bryant P.J.,
RA      Small K., Forsman-Semb K.;
RT      "Genomic organization of human DLG4, the gene encoding post-synaptic
RT      density 95 (PSD95).";
RL      J. Neurochem. 73:0-0(1999).
RN      [4]
RP      SEQUENCE OF 81-401 FROM N.A.
RC      TISSUE=Brain;
RC      Breman J.E., Bredt D.S., Parkinson J.F., Manzana W.P., McClary J.A.;
RL      Submitted (AUG-1996) to the EMBL/Genbank/DBJ databases.
CC      -1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR
CC      SUBUNITS. MAY BE INVOLVED IN SYNAPTOSIS.
CC      -1- SUBCELLULAR LOCATION: CONCENTRATED AT SYNAPTIC JUNCTIONS PRIMARILY
CC      ON THE PRESYNAPTIC SIDE (WAS ORIGINALLY THOUGHT TO BE
CC      POSTSYNAPTIC).
CC      -1- TISSUE SPECIFICITY: PRESYNAPTIC DENSITY FRACTION OF BRAIN.
CC      -1- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
CC      -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC      -1- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
CC      -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL; U83192; AAC52113.1; -
DR      EMBL; AF156495; RAD56173.1; -
DR      EMBL; U68138; AAB07736.1; -
DR      HSPB; Q12959; 1PDR.
DR      MIM: 602887; -
DR      InterPro; IPR000619; -
DR      InterPro; IPR001452; -
DR      InterPro; IPR001478; -
DR      Pfam; PF00625; Guanylate_kin; 1.
DR      Pfam; PF00595; PDZ; 3.
DR      Pfam; PF00018; SH3; 1.
DR      PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
DR      PROSITE; PS00502; GUANYLATE_KINASE_2; 1.
DR      PROSITE; PS50106; PDZ; 3.
DR      PROSITE; PS50002; SH3; 1.
KW      SH3 domain; Repeat.

```

```

FT      DOMAIN      108      194      PDZ 1.
FT      DOMAIN      203      289      PDZ 2.
FT      DOMAIN      356      436      PDZ 3.
FT      DOMAIN      471      541      SH3.
FT      DOMAIN      577      767      GUANYLATE KINASE.
FT      CONFLICT      46      46      E -> V (IN REF. 3).
FT      CONFLICT      81      83      VIV -> EFR (IN REF. 4).
FT      CONFLICT      399      401      GDO -> AGI (IN REF. 4).
SQ      SEQUENCE      767 AA; 85429 MW; BE1019159E65B2D8 CRC64;

Query Match      5.5%; Score 384; DB 1; Length 767;
Best Local Similarity      25.8%; Pred. No. 5.3e-13;
Matches 131; Conservative 76; Mismatches 153; Indels 148; Gaps 18;

QY      886      NSQSILMPQSGSP-----EPESIRNTSRSSSTPAIRASDPATCP-----IIPGCTTIE--- 934
DB      50      SESQKRYODEDPTPLPHESPAPLIPNOA-NSPPYVTFVTDLEAPGYELQVNGTEGEMEYEE 108
QY      935      --ISKRTGLGISVGSPT-LLG---AFILHVEYEGAAKCDGRMLMAGDQILEVNGIDL 988
DB      109      ITLERKNSGLGFSIAGTDNPHIGDDPSITFTKITPGGAADGRRLVNDLSILEVNEVDV 168
QY      989      KRATHDEAINVLKQTPQRYRLTLRDAPYKEEYCDTLTIELQKKPKGLGISVSK-- 1046
DB      169      REYTHSAAVEALKENGSIVRLVYMRKPPAEK-----VMEIKLIKRP-KGLGFSINGVG 222
QY      1047      ---RNDTGVFSDIVYKGIADPDGRLLGDDQILLVNGEDVRRNASQEAVALLKSLGTV 1102
DB      223      NQHIPGDNISIVYKILLEGGAHRKDRLQIGDKILAVNSVLEGVMEHDAVVALKNTYDVV 282
QY      1103      TLEVGRIRKAGPFSERKPSQTSQVSEGLSFPFLSGSSTSESSSKKNALASEIQ 1162
DB      283      YLKVA-----KPS-NATLSD-----STAPPITTSYSHLDNEISHS----- 319
QY      1163      LRTVENKKGPTDSLGISIAGVGSPGLGDVPIFAMHPGVAAQOTKRLVGRIVTTCGT 1222
DB      320      -----YLGCTYPP----- 327
QY      1223      STEGMYHTQAVNLKNAASISIEQVYAGDVSVYTGHHQEPASSLSFTGLSTSTFQDD 1282
DB      328      ---AMTPT-----SPRAYSVPVKDLGLED 349
QY      1283      LGGPPQITLERGPGGLGSIVGSGPHGDLPIYKTVFAKGAASEGRKRGDQITA 1342
DB      350      I-PRERRIVIHRSSTGLGFNTVG---EDGE-GIFSTIAGPADLSGELKKGQJLS 404
QY      1343      VNGQSLEGVTHEEVAVALIKRTGVTIM 1370
DB      405      VNGVDLRNASHDAQALALNNAQTVTII 432

RESULT 8
ID      DLG4_RAT      STANDARD:      PRT:      724 AA.
AC      P31016; P97631;
DT      01-JUL-1993 (Rel. 26; Created)
DT      01-JUL-1993 (Rel. 26; Last sequence update)
DT      01-OCT-2000 (Rel. 40; Last annotation update)
DE      PRESYNAPTIC DENSITY PROTEIN 95 (PSD-95) (PRESYNAPTIC PROTEIN SAP90)
DE      (SYNAPSE-ASSOCIATED PROTEIN 90) (DISCS, LARGE HOMOLOG 4).
GN      DLG4 OR PSD95.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
OX      NCBI_Taxid=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RC      MEDLINE=93040233; PubMed=1419001;
RA      Cho K.-O., Hunt C.A., Kennedy M.B.;
RT      "The rat brain postsynaptic density fraction contains a homolog of
RT      the Drosophila discs-large tumor suppressor protein.";

```

```

RL  Neuron 9:929--942(1992).
RN
RP  SEQUENCE FROM N.A.
RC  STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RX  MEDLINE=93186749; PubMed=7680343;
RA  Kistner U., Wenzel B.M., Voh R.W., Cases-Langhoff C., Garner A.M.,
RT  Appelbauer U., Voss B., Gundelfinger E.D., Garner C.C.;
RT  "SAP90, a rat presynaptic protein related to the product of the
RL  Drosophila tumor suppressor gene dlg-A.";
RN  J. Biol. Chem. 268:4580-4583(1993).
RN
RN  [3]
RP  SEQUENCE OF 566-625 FROM N.A.
RC  STRAIN=WISTAR KYOTO; TISSUE=Vascular smooth muscle;
RA  Adams L.D., Werny I., Schwartz S.M.;
RL  Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
RN
RN  [4]
RP  X-RAY CRYSTALLOGRAPHY (1.82 ANGSTROMS) OF 302-402.
RX  MEDLINE=96270509; PubMed=8674113;
RA  Doyle D.A., Lee A., Lewis J., Kim E., Sheng M., Mackinnon R.;
RT  "Crystal structures of a complexed and peptide-free membrane protein-
RT  binding domain: molecular basis of peptide recognition by PDZ.";
RL  Cell 85:1067-1076(1996).
CC  -1 FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR
CC  SUBUNTS. MAY BE INVOLVED IN SYNAPTOGENESIS.
CC  -1 SUBCELLULAR LOCATION: CONCENTRATED AT SYNAPTIC JUNCTIONS PRIMARILY
CC  ON THE PRESYNAPTIC SIDE (WAS ORIGINALLY THOUGHT TO BE
CC  POSTSYNAPTIC).
CC  -1 TISSUE SPECIFICITY: PRESYNAPTIC DENSITY FRACTION OF BRAIN.
CC  -1 SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
CC  -1 SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC  -1 SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
CC  -1 SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC
CC  THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
CC  BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -
CC  THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON
CC  USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
CC  COMMERCIALIZED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
CC  ENTITIES REQUIRES A LICENSE AGREEMENT (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC
DR  EMBL; M96853; AAA41971.1; -
DR  EMBL; X66474; CAA47103.1; -
DR  EMBL; U77090; AAB38270.1; -
DR  PIR; S26407; S26407.
DR  PIR; JH0800; JH0800.
DR  PDB; 1BE9; 21-OCT-98.
DR  PDB; 1BFE; 21-OCT-98.
DR  InterPro; IPR000619; -
DR  InterPro; IPR001478; -
DR  InterPro; IPR001452; -
DR  Pfam; PF00625; Guanylate_kin; 1.
DR  Pfam; PF00595; PDZ; 3.
DR  Pfam; PF00018; SH3; 1.
DR  PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
DR  PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
DR  PROSITE; PS50106; PDZ; 3.
DR  PROSITE; PS50002; SH3; 1.
KM  SH3 domain; Repeat: 3D-structure.
FT  DOMAIN 65 151 PDZ 1.
FT  DOMAIN 160 246 PDZ 2.
FT  DOMAIN 313 393 PDZ 3.
FT  DOMAIN 428 498 SH3.
FT  DOMAIN 534 724 GUANYLATE KINASE.
FT  CONFLICT 61 61 M -> L (IN REF. 2).
FT  CONFLICT 78 78 S -> T (IN REF. 2).
FT  CONFLICT 177 182 G -> G (IN REF. 2).
FT  CONFLICT 200 200 A -> G (IN REF. 2).
FT  CONFLICT 254 254 S -> T (IN REF. 2).
FT  CONFLICT 539 555 IGLPTKQRANDLLSEF -> ISLDPKTYVTMLFSSS
FT  (IN REF. 2).
FT  CONFLICT 623 625 GKH -> RDO (IN REF. 3).
FT  SEQUENCE 724 AA; 80465 MW; 7922D4EB0F9AD85 CCK64;

```

Query Match	Similarity	Score	DB 1	Length	724
Best Local Similarity	26.2%	Pred. No. 5.5e-13			
Matches 128:	Conservative	74	Mismatches	145	Indels 144; Gaps 17
QY	901	PEINSTRSSSTPAIFASDPTCP----	IIPECEETIE----	ISKRTGLSLGVGSD	951
Db	26	PAHLRNOA--NSPPVIVNVDLEAPYEIV	OVNTEGEMEETTLRGNSGLFSGIS	IND	84
QY	952	T-LIG--AFITHEVEGSAACKDGR	ILVNGIDILKATNDEINVLROTPORV	1007	
Db	85	NPHIDDDSIPTFKIIPGGAAD	ORLKVNSDILVEYNEVDNEVTHSA	VEVLEKAGSIV	144
QY	1008	RLTLVRDEAPRKKEEVCSTLT	LELQKKRKGISLVK----	RNDTGVFSDIVAGG	1061
Db	145	RLTYVRKKRPAPKK-----	VMETIKIKRP-KGLGFSIAGVGNQ	HPGNSIVYTKIIEGG	198
QY	1062	IADPRGLIGDQIILLVNGED	VRNMSQEVAAALLCSLGTYL	LEVGRIKAGFHSRRPS	1121
Db	199	AAHKGRIGIIGKILIAVVS	GLEDVMDDAVALKNTYDVYLKVA	-----RPS	247
QY	1122	QTSQVSEGLSSEFPRLSGS	STSESLSSKKNALASEIOLGRT	VEMKKGPDSIGISIA	1181
Db	248	-NATLSD-----SYAPRD	ITTSYSQHLDEISMSS-----	-----	276
QY	1182	GGVSPCLDDVPIFFIMAH	PTGYAOTOKLRNGDRITVTCIG	STESMTPTQAVNLKMSG	1241
Db	277	-----YLDTPPT-----	-----MTPP-----	-----	289
QY	1242	SIEMQVAVGADVVTGHN	HOERASSLSFTGLTSTISIFOD	LDGRPOCKSYTLERPDGLG	1301
Db	290	-----SPKRYSPAKDLLEED	-PREPRRIYVHNRSSTGLG	324	
QY	1302	FSIVGGYSSPHGDLPIYK	YVFAKGAASEDGRKRGDQII	AVNGSLEGVTHEEVAALIK	1361
Db	335	FNIVGG---EDGE-GIF	SFILAGPADLSCGLARKGDQI	ILSYNGVDLRNASHEDQAALIK	380
QY	1362	RTKGVTLM	1370		
Db	381	NAGQTVTII	389		
RESULT	9				
DLG4_MOUSE		STANDARD:	PRT:	724 AA.	
ID	DLG4_MOUSE				
AC	Q62108:				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	01-OCT-2000 (Rel. 40, Last annotation update)				
DE	PRESYNAPTIC DENSITY PROTEIN 95 (PSD-95) (PRESYNAPTIC PROTEIN SAP90)				
DE	(SYNAPSE-ASSOCIATED PROTEIN 90) (DISCS, LARGE HOMOLOG 4).				
GN	DLG4 OR DLG4 OR PSD95.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eulambda; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=DBA/2; TISSUE=Brain;				
RA	Kohmura N., Yagi T.;				
RL	Submitted (May-1995) to the EMBL/GenBank/DBJ databases.				
CC	-1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR SUBUNITS. MAY BE INVOLVED IN SYNAPTOGENESIS (BY SIMILARITY).				
CC	-1- SUBCELLULAR LOCATION: CONCENTRATED AT SYNAPTIC JUNCTIONS PRIMARILY ON THE PRESYNAPTIC SIDE (BY SIMILARITY).				
CC	-1- SIMILARITY: CONTAINS 3 PDZ/DRH DOMAINS.				
CC	-1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.				
CC	-1- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.				
CC	-1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.				
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation -				

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).

DR EMBL: D50621; BAA09297.1; -
 DR HSP: O12959; 1PDR.
 DR MGP: MGI:127959; DlgH4.
 DR InterPro: IPR000619; -
 DR InterPro: IPR001452; -
 DR InterPro: IPR001478; -
 DR Pfam: PF00625; Guanylate_kin; 1.
 DR Pfam: PF00595; PDZ; 3.
 DR Pfam: PF00018; SH3; 1.
 DR PROSITE: PS00856; GUANYLATE_KINASE_1; 1.
 DR PROSITE: PS50052; GUANYLATE_KINASE_2; 1.
 DR PROSITE: PS50106; PDZ; 3.
 DR PROSITE: PS50002; SH3; 1.
 DR SH3 domain; Repeat.
 FT DOMAIN 65 151 PDZ 1.
 FT DOMAIN 160 246 PDZ 2.
 FT DOMAIN 313 393 PDZ 3.
 FT DOMAIN 428 498 SH3.
 FT DOMAIN 534 724 GUANYLATE_KINASE.
 SO SEQUENCE 724 AA; 80472 MW; 7EFEC99E1FF90BA CRC64;

Query Match 5.5%; Score 382; DB 1; Length 724;
 Best Local Similarity 26.2%; Pred. No. 6.2e-13;
 Matches 128; Conservative 74; Mismatches 143; Indels 144; Gaps 17;

QY 901 PESIRNTSSSPATPASPATP-----IIPCETTF-----ISKRGGLGSLVSGSD 951
 DB PAHLPMQA-NSPPIVNTTLEAPGELOYNGEGEMEYEETTLERNGSLGSIAGTD 84
 QY 952 T-LG---AFITHEVEEGACGRLMAGDOILEVNGIDLRKATDEALINVLKOTPOVY 1007
 DB 85 NHIGDPSIFITIKITPGMAADGRLRVNDSTLPEVNEVDREVTISAVALKEAGSIV 144
 QY 1008 RLLVDEAPYKEEYVCDTLTLELQKPKGLSLVSK-----RNDGVFVSDIVKGG 1061
 DB 145 RLYVMRRKPPAEK-----ITIKLKGP-KGLGFSIAGGVGNHIGDINSIYVTKIIEGG 198
 QY 1062 IADPGRQLQSGOQIILVNEEDVNAQSEVAALLKSLGTVLLEVRIKAGPRHSRRS 1121
 DB 199 AAHKDRLOIGKILAVNSVGLVEDVNAVALKNTYVYLKVA-----KPS 247
 QY 1122 QTSOVSEGLSFTPLSGSTSESSSKNALASEIQLRTVEMKKGPDSIGISIA 1181
 DB 248 -NAVLSLSD---STAPPDITTSYSHLDNELSHS----- 276
 QY 1182 GGVGSPGLGVPIFIAMHPTGVAAGQOKLAVGDRIVTICGTSTEGMTHTQAVNLLKNASG 1241
 DB 277 -----YLGTDYPT-----AATPT----- 289
 QY 1242 SIEMOVVAGDVSVYTGHHOEPASSLSFTGLSTSIIFODDLGPPCKSTTLERGDGIC 1301
 DB 290 -----SPRRYSVPAKDLLEGEDI-PRPRRIVYHRSIGTGIG 324
 QY 1302 FSTVGGSPHGDLPYVTVFKGAASEDRGLRGDOITIAVNGSLLEGVTHEEVAAILK 1361
 DB 325 FNVVGG---EDGE-GIFISFIILAGPADLSGELRKGDQLISVNGVLDLRNASHQAAIALK 380
 QY 1362 RTKGVITLM 1370
 DB 381 NAGQVTVII 389

RESULT 10
 DGLI_DROME STANDARD; PRT; 960 AA.
 ID DGLI_DROME
 AC P31007;

DR 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DE 30-MAY-2000 (Rel. 39, Last annotation update)
 DE LETHAL(1)DISCS LARGE-1 TUMOR SUPPRESSOR PROTEIN.
 GN DGLI.

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriidae; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91330294; PubMed=1651169;
 RA Woods D.F., Bryant P.J.;
 RT "The disc-large tumor suppressor gene of Drosophila encodes a
 RT guanylate kinase homolog localized at septate junctions.";
 RL Cell 66:451-464(1991).

-1- FUNCTION: LOSS OF THIS PROTEIN CAUSES LARGE IMAGINAL DISKS BY
 ALLOWING NEOPLASTIC OVERGROWTH. THE PRESENCE OF A GUANYLATE KINASE
 DOMAIN SUGGESTS INVOLVEMENT IN ADEQUATE CELLULAR PROLIFERATION. THE DGL
 AS SIGNAL TRANSDUCTION TO CONTROL CELLULAR PROLIFERATION. THE DGL
 PROTEIN PLAYS A CRITICAL ROLE AT SEPTATE JUNCTIONS IN CELLULAR
 GROWTH CONTROL DURING LARVAL DEVELOPMENT.

-1- SUBCELLULAR LOCATION: CYTOSKELETON-ASSOCIATED. LOCATED AT THE
 CYTOPLASMIC FACE OF THE MEMBRANE IN THE CELLULAR BLASTODERM AND
 BECOMES ASSOCIATED WITH SEPTATE JUNCTIONS WHICH BEGIN TO FORM
 BETWEEN EPITHELIAL CELLS AT THE TIME OF DORSAL CLOSURE. IN ADULT
 FLIES, LOCATED AT THE APICAL-LATERAL MEMBRANE BOUNDARY OF
 EPITHELIAL CELLS.

-1- ALTERNATIVE PRODUCTS: LEADS TO DIFFERENT FORMS OF THE PROTEIN;
 THE SEQUENCE SHOWN REFERS TO THE DGL-A PROTEIN, THAT IS MOST
 ABUNDANTLY EXPRESSED.

-1- TISSUE SPECIFICITY: BROAD SPECIFICITY. ABUNDANTLY EXPRESSED IN
 EPITHELIAL TISSUES. DEVELOPMENTAL STAGE DEPENDENT EXPRESSION.
 CC -1- SIMILARITY: CONTAINS 3 PDZ/DRH DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).

DR EMBL: M73529; AAA28468.1; -
 DR PIR: A39651; A39651.
 DR HSP: O12959; 1PDR.
 DR FlyBase: FBgn0001624; dgl1.
 DR InterPro: IPR000619; -
 DR InterPro: IPR001452; -
 DR InterPro: IPR001478; -
 DR Pfam: PF00625; Guanylate_kin; 1.
 DR Pfam: PF00595; PDZ; 3.
 DR Pfam: PF00018; SH3; 1.
 DR PROSITE: PS00856; GUANYLATE_KINASE_1; 1.
 DR PROSITE: PS50052; GUANYLATE_KINASE_2; 1.
 DR PROSITE: PS50106; PDZ; 3.
 DR PROSITE: PS50002; SH3; 1.
 KW Transducer; SH3 domain; Alternative splicing; Repeat.
 FT DOMAIN 40 126 PDZ 1.
 FT DOMAIN 154 244 PDZ 2.
 FT DOMAIN 486 566 PDZ 3.
 FT DOMAIN 600 670 SH3.
 FT DOMAIN 768 960 GUANYLATE_KINASE.
 SO SEQUENCE 960 AA; 102468 MW; BF87A4262F1B6AD5 CRC64;

Query Match 5.4%; Score 379; DB 1; Length 960;
 Best Local Similarity 24.8%; Pred. No. 1.4e-12;
 Matches 136; Conservative 85; Mismatches 190; Indels 138; Gaps 19;

```

QY 933 IEISKRTGLSLVSGSDT-LLG---AFIHEVEEGAACKGRIMAGDQILEVNGIDL 988
DB 41 IQLEHNGSLGSLVSGSDTNDPHIGDTISYITIKLSGGAADGSLINDIVSNVDSV 100
QY 989 KKAHDEAIINVLKROTPOQRVLTLYRDE-----APKKEEVCDTTLTLEQ-KKP 1035
DB 101 VDVPHASAVDLKAKGNVYKLVKRRKCTATTPAAGSAAGRDASAGPKYIEIDLK 160
QY 1036 GKGGLSLVSGK-----RNDGVFVSDIVKGIADPDGRLLQDQIILV-NG-EDVRN 1085
DB 161 GKGGLSLVSGKIGNQHITGDMGIVTKLTDGRQVDRSLTIGDLIVKRTNGSEKLN 220
QY 1086 ASQEAVALLCSCIGTTLTLEVGRKAKGPFHSERRPSQTSQVSEGLSFTPELSSSTSE 1145
DB 221 VTHELAVALTLKSTFDKVLTIK-----TQHLTTSASGGGGGLS-----SQQLSQ 267
QY 1146 S-----LESSKNALASEIQLRTVEMK-CPDLSIGISIAGVSGPLDVPFTIMMHP 1201
DB 268 SOSQALATISQSOVHQOQHATPMVNSOSTGALNSMGQTV---VDS-SP-SDPAAAAVAA 322
QY 1202 GVAATOKLRYDRIYTCIGTSTEGMTHTQAVNLKNAAGSIEMQVAVAGDVSV----- 1255
DB 323 ANASASAVIASN- -TISNTVTYVATATAS---NDSSKLPBGLANSISISNSN 377
QY 1256 -----VTGHIQEPASSLSFTGLTSTST-----FODDLGPP----- 1286
DB 378 SNSNNNNNNNNNNSSSSSTTATVAAATPTAASAAAAAASPANFYNASMPALPV 437
QY 1287 -----QCSITLERPDGL 1301
DB 438 ESNQTNRSQSPQRPQSPKRASTNVLAAPPPTPRAVSTEDTREPPTTIQKPGGLG 497
QY 1302 FSIYGVGSPGHDPIYKYTFKAGASEDRKRGDQIIAVNGSLGVTHEEVAALIK 1361
DB 498 FNIYGG-----EDGGGIYVFLIAGSPADLSELRKGDLLSVNNVNLTHAHEEAQALK 553
QY 1362 RTKGTVTLM 1370
DB 554 TSGGVVTL 562

```

RESULT 11
DLGL_RAT
ID DLGL_RAT STANDARD: PRT: 911 AA.
AC Q62696;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PRESYNAPTIC PROTEIN SAP97 (SYNAPSE-ASSOCIATED PROTEIN 97) (DISCS, LARGE HOMOLOG 1).
GN DLGL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RX MEDLINE=95198112; PubMed=7891172;
RA Mueller B.M., Kistner U., Voh R.W., Cases-Langhoff C., Becker B., Gundelfinger E.D., Garner C.C.;
RT "Molecular characterization and spatial distribution of SAP97, a novel presynaptic protein homologous to SAP90 and the Drosophila discs-large tumor suppressor protein."
RL J. Neurosci. 15:2354-2366(1995).
CC -1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR SUBUNITS. MAY BE INVOLVED IN SYNAPTOGENESIS. MAY PLAY A ROLE IN CELL ADHESION.
CC -1- SUBCELLULAR LOCATION: IN CULTURED T84 CELLS, IT IS RESTRICTED TO THE CYTOPLASMIC SURFACE OF THE PLASMA MEMBRANES BETWEEN ADJACENT CELLS, BUT NOT AT THE EDGES OF CELLS LACKING CELL CONTACT.
CC -1- TISSUE SPECIFICITY: PRESYNAPTIC NERVE TERMINI OF EXCITATORY

```

CC SYNAPSES. IN OTHER BRAIN REGIONS, SAP97 IS FOUND IN AND ALONG
CC BUNDLES OF UNMYELINATED AXONS. SAP97 IS NOT RESTRICTED TO THE CNS,
CC BUT IS ALSO PRESENT AT THE BASAL LATERAL MEMBRANE BETWEEN A
CC VARIETY OF EPITHELIAL CELLS.
CC -1- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC -----
DR EMBL: U14950; AAA79976.1; -.
DR HSSP: Q12959; 1PDR.
DR InterPro: IPR000619; -.
DR InterPro: IPR001452; -.
DR InterPro: IPR001478; -.
DR Pfam: PF006925; Guanylate_kin. 1.
DR Pfam: PF00595; PDZ; 3.
DR Pfam: PF00018; SH3; 1.
DR PROSITE: PS00856; GUANYLATE_KINASE_1; 1.
DR PROSITE: PS50052; GUANYLATE_KINASE_2; 1.
DR PROSITE: PS50106; PDZ; 3.
DR PROSITE: PS50002; SH3; 1.
DR SH3 domain; Repeat.
KW SH3 domain; Repeat.
FT DOMAIN 224 310 PDZ 1.
FT DOMAIN 318 404 PDZ 2.
FT DOMAIN 465 545 PDZ 3.
FT DOMAIN 580 650 SH3.
FT DOMAIN 721 911 GUANYLATE_KINASE.
FT DOMAIN 527 530 POLY-ALA.
SQ SEQUENCE 911 AA; 100570 MW; 18CEBD31DDCAF8B CRC64;

```

Query Match 5.48; Score 378.5; DB 1; Length 911;
Best Local Similarity 23.38; Pred. No. 1.3e-12;
Matches 151; Conservative 91; Mismatches 201; Indels 205; Gaps 21;

```

QY 774 AAYDLSPKNOVHLEPK-----DQGLGIAISEDTLSGVII----- 811
DB 48 ALIDHGEFYEVLTDNPKCVHDKCEPVQPNPWESSLSAAVTSLSLPGASPVYK 107
QY 812 -----KSLTEHGAATGRKLVGDIIVD-----EIVVGPIEFISLTKAKTV 859
DB 108 YRYODEEVLPSERISPOVPEVNLGELVHSEKSLSEIENHGVFVSHSHPIKPT- 164
QY 860 KLTTHAENPDQAVPSAAGSAGSEKNSOSILAMPQSGPPESTIRMSRSSTPAIFASD 919
DB 165 -----AVPPSSPIY-----VTPALPVP---AESPPVLPSPQANPPVLVNT 204
QY 920 PA--TCPILPCETITE-----ISKRTGLSLVSGSDT-LLG---AFIHEVEEGA 968
DB 205 DSELPYVNGTDAVDEEETLERGNSGLGSLAGTDNDPHIGDSSIFITKIITGGA 264
QY 969 CKDGLMAGDQILEVNGIDLRKATHDEAIINVLKROTPOQRVLTLYRDEAPKKEEVCDTLT 1028
DB 265 AQDRLRLVNDILRVNEDVADVTHSKAVELKAGSVLRVLYVRRKAFRRKHE----- 318
QY 1029 IELKKKKGKGLSLVSGK-----RNDGVFVSDIVKGIADPDGRLLQDQIILVNGED 1082
DB 319 IKLKGP-KGLGFSIAGVGNQHPDGNSTVYTIIEGAHKGKLGKQIKLAVNSVC 377
QY 1083 VRNASQEAVALLCSCIGTTLTLEVGRKAKGPFHSERRPSQTSQVSEGLSFTPELSSGS 1142
DB 378 LEEVTHEAVTALNTDQVLYLKA-----PISMYINDG-----YAPPDITNS 421
QY 1143 TSELESSKNALASEIQLRTVEMKGPDSIGISIAGVSGPLDVPFTIMMHP 1202

```



```

Db 422 SSQSVDNHVSPTS-----SYLGQTPASPARYSPTS 450
QY 1203 VAAQIOKLVKGVDRIVITCGSTIEGMDTHQAVNLKNAAGSIEMQVAVAGDVSVYTHHDE 1262
Db 451 KAV-----LGDEIT-----RD 462
QY 1263 PASSSLFTGLTSTSFQDGLGPPCKSTLREPGDLSFYGVGSPHGLPIYVKRY 1322
Db 463 P-----RKVYLHNGSTGLGFNIYG---EDGE-GITISFT 493
QY 1323 FAKGAASEDGRKRGDQIIAIVNGQSLEGVTHEEVAAILKRTKGTVTLM 1370
Db 494 LAGGPADLSGELKKGRITISVNSVDLRAASHEDQAAALKNAGQAVTIV 541

RESULT 12
ID LG1_HUMAN STANDARD; PRT; 904 AA.
AC 012959; 012958;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PRESYNAPTIC PROTEIN SAP97 (SYNAPSE-ASSOCIATED PROTEIN 97) (DISCS,
DE LARGE HOMOLOG 1).
GN DLG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95024052; PubMed=7937897;
RA Lue R.A., Marfatia S.M., Branton D., Chishti A.H.,
RT "Cloning and characterization of hdlg1, the human homologue of the
RT Drosophila discs large tumor suppressor binds to protein 4.1.";
RT Proc. Natl. Acad. Sci. U.S.A. 91:9818-9822(1994).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 460-555.
RX MEDLINE=96338231; PubMed=8757139;
RA Cabral J.H., Petosa C., Sutcliffe M.J., Raza S., Byron O., Poy F.,
RA Marfatia S.M., Chishti A.H., Liddington R.C.;
RT "Crystal structure of a PDZ domain.";
RT Nature 382:649-652(1996).
CC -!- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR
CC SUBUNITS. ASSOCIATES WITH PROTEIN 4.1.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 (SHOWN HERE) AND 2; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: CONTAINS 3 PDZ/DRH DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -!- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U13897; AAA50599.1; -
CC EMBL: U13896; AAA50598.1; -
CC PDB: 1PDR; 23-JUL-97.
CC MIM: 601014; -
CC InterPro: IPR000619; -
CC InterPro: IPR001452; -
CC InterPro: IPR001478; -
CC Pfam: PF00625; Guanylate_kin; 1.
CC Pfam: PF00059; PDZ; 3.
CC Pfam: PF00018; SH3; 1.
CC PROSITE: PS00856; GUANYLATE_KINASE_1; 1.
CC PROSITE: PS00552; GUANYLATE_KINASE_2; 1.
CC PROSITE: PS0106; PDZ; 3.

```

```

DR PROSITE: PS00002; SH3; 1.
KW SH3 domain; Repeat; Alternative splicing; 3D-structure.
FT DOMAIN 224 310 PDZ 1.
FT DOMAIN 319 405 PDZ 2.
FT DOMAIN 466 546 PDZ 3.
FT DOMAIN 581 651 SH3.
FT DOMAIN 714 904 GUANYLATE KINASE.
FT VARSPIC 669 680 ETPDDKSGSLK -> QSNDRKRLKRLSPKPPFYKNDQS
FT ETPDDKSGSLK (IN ISOFORM 2).
SQ SEQUENCE 904 AA; 100354 MM; B78798DBB020D4 CRC64;

Query Match 5 4%; Score 377; DB 1; Length 904;
Best Local Similarity 29.9%; Pred. No. 1,6e-12;
Matches 119; Conservative 58; Mismatches 129; Indels 92; Gaps 14;

QY 1002 QTPQRYRLTYRDEAPYKEEVECDTLIEQKPKGKLSYGRK-----NDTGVS 1055
Db 208 EPTTYVNGT-----DADYEVEI-----TLE---RNGSLGSLNGSTGNDPHIGDSSIFIT 256
QY 1056 DIVKGGIADPDGRLIGDQIILVNGEDVNASQEAVALKLSGLGVTVLEVRKAKPRH 1115
Db 257 KITTGAAADGRLRYNDCILQVNEVDVDRVTSKAVEALKKAGSIYRLYVKR----- 309
QY 1116 SERPSQTSQVSEGSLSSTFFPLSGSTSSLESSKKNALAEIOGLRTVEKKKPTDS 1175
Db 310 --RKP-----VSE-----KIMEIKLIKGP-KG 328
QY 1176 LGISLGVGSP--LGDVPIFIAHPTGYAAQOTKRLVGRDRIYITCGSTEGMTQAV 1233
Db 329 LGFSIAGVGNGNHPDGNSTYVTKIEGGAHNDGKLQIGDKLAVNNVCLEEVTHEAV 368
QY 1234 NLKNAAGSIEMQVAVAGDVSV-----VTGHQDEPAS---SSLSFTGLT----- 1274
Db 389 TALKNTSDFYLYKVAKPTSMYNDGYAPPDITNSSQPVNDHVSPPSLGQTPASPARYS 448
QY 1275 --STSIQDGLGPPCKSTLREPGDLSFYGVGSPHGLPIYVKRYFAKGAASEDG 1332
Db 449 PVSKAVALGDEITREPRKRVYLHNGSTGLGFNIYG---EDGE-GITISFTLAGGPADLSG 504
QY 1333 RLKRGDQIIAIVNGQSLEGVTHEEVAAILKRTKGTVTLM 1370
Db 505 ELKKGRITISVNSVDLRAASHEDQAAALKNAGQAVTIV 542

RESULT 13
ID IL16_HUMAN STANDARD; PRT; 631 AA.
AC 014005; 016435; 090P18;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE INTERLEUKIN-16 PRECURSOR (IL-16) (LYMPHOCYTE CHEMOTACTANT FACTOR)
DE (LCP).
GN IL16.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=peripheral blood;
RA Kornfeld H.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99145586; PubMed=9990060;
RA Bannert N., Avots A., Baler M., Serfling E., Kurth R.;
RT "GA-binding protein factors, in concert with the coactivator CREB
RT binding protein/p300, control the induction of the interleukin 16
RT promoter in T lymphocytes.";
RT Proc. Natl. Acad. Sci. U.S.A. 96:1541-1546(1999).
RN [3]

```



```

Oy      1277 SI-PDOLLGPPQCKSITLERGP-----1297
Db      343 PVKADHDPTVEVETATYERNEKOTPSLPEPKRPVYAQNOMWITYLSHLMVSYLIQLMK 402
Oy      1298 -----DELGSIVGGYGSPHGDLPYVKTYFAFKASEDSRLKRGDOI 1341
Db      403 MGLFARSKKLKYFRKGDSVGLRAG-----NDVGIFPAVGLDESPAKREG-LECGDIIL 456
Oy      1342 AVNGOSLEGVTHEAVAIL 1360
Db      457 RYNVDFTNIIREEVAFL 475

RESULT 15
Z03_HUMAN
ID Z03_HUMAN STANDARD; PRT; 933 AA.
AC 095049;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DEF TIGHT JUNCTION PROTEIN ZO-3 (ZONA OCCUDENS 3 PROTEIN) (TIGHT JUNCTION
DE PROTEIN 3).
GN TPJ3 OR ZO3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Vismathan V.,
RA Burkhardt-Schultz K., Brower A., Gordon L., Dias J., Ramirez M.,
RA Stillwagen S., Phan H., Velasco N., De La, Regala W., Terry A.,
RA Ganes J., Dangsan L., Erlar A., Christensen M., Georgescu A.,
RA Avila J., Liu S., Attix C., Andreise T., Frankheim M.,
RA Amico-Keller G., Coetfield J., Duarte S., Lucas S., Bruce R.,
RA Thomas P., Quan G., Kronmiller B., Arellano A., Sanders C., Ow D.,
RA Noian M., Trong S., Kobayashi A., Olsen A.S., Carraro A.V.;
RT "Sequence analysis of a 3.5 kb contig in human 19p13.3 containing a
RT serine protease gene cluster";
RL Submitted (NOV-1998) to the EMBL/Genbank/DDBJ databases.
CC -1 SUBUNIT: INTERACTS WITH OCCUDLIN, CLAUDINS AND ZO-1 (BY
CC SIMILARITY).
CC -1 SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
CC -1 SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1 SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
CC -1 SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AC005954; AAC72274.1; ALT_INIT.
DR HSSP: P31016; 1BFE.
DR InterPro: IPR000619; -.
DR InterPro: IPR001452; -.
DR InterPro: IPR001478; -.
DR Pfam: PF00595; PDZ_3.
DR DR PROSITE: PS00625; Guanylate_kin; 1.
DR PROSITE: PS00856; GUANYLATE_KINASE_1; FALSE_NEG.
DR PROSITE: PS50052; GUANYLATE_KINASE_2; 1.
DR PROSITE: PS50002; SH3; 1.
DR PROSITE: PS50106; PDZ; 3.
DR KW Tight junction; SH3 domain; Repeat; Membrane.
DR FT DOMAIN 11 93 PDZ 1.
DR FT DOMAIN 195 272 PDZ 2.
DR FT DOMAIN 394 460 PDZ 3.
DR FT* DOMAIN 489 563 SH3.
DR FT DOMAIN 675 775 GUANYLATE_KINASE.

```

SO SEQUENCE 933 AA; 102800 MW; 680298CFD0615B47 CRC64;

Query Match 3.5%; Score 242; DB 1; Length 933;

Best Local Similarity 22.8%; Pred. No. 2.3e-05;

Matches 121; Conservative 62; Mismatches 155; Indels 192; Gaps 20;

```
QY 932 TLEISKG-RTGLSLIVGSDTLGAFTHIEYEEGACAKDGLWAGDQILEVNGIDLK 990
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 11 TATLSKDPRRRGIAISGDRPGSGMVSVDVPGGPA--EGRLQGDHIVAVNGVSMEN 68
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 991 ATHDEAIVLR-----OTFORVRLTLR-----DEAPYKEEV----- 1023
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 69 ATSAFAIQILKTCFMANITVKKPRRIHLPATKASPSGRODDEDDGQPRVEEVQGR 128
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1024 -----CDTLT-----EL 1031
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 129 GYDGDSSSGSGRSDERRRPRPGRRAGSHGRSPGGSGSEANGLALVSGFKRLPRODV 188
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1032 QKKPGKGLSLIVGKRD-----TGVPVSDIVKGIADPDGRLIQDQILLVNGED 1082
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 189 QMKPYK-----SVLVKRRDSEEGVVLGSIQIFIKHTDGLAARHRLQEGDLIIQINGVS 244
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1083 VRNASQEAVALKCSLCTVLEVR----- 1108
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 245 SONSLANDTRRRLIEKSEKLSILVLRDGOFLVNIIPAVSDSDSSPLEEGVTMADEMSSP 304
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1109 -----IKAGPFHSERRP--SOTSQV-----SEGLSFPPLSGS 1141
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 305 PADISDLASELSQAPPSHIIPPPRHAQNSPEASQSDSPVESRRLRRESSVDSRTISEPDE 364
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1142 STSELESSES-----KKNALASEIOG---LRYEMKKGPTDSLIGISIAGVGSPLDVP 1192
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 365 QRSELPRESSYDIYVPSQSMEDRGYSPTRVVRFELK--KSIQRLAGG-----NDVG 417
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1193 IEFIMMHPTGVAAGTQKLRVGDRTITIGTSTEGHTQAVNLKNAGSIEMQVYAG-- 1250
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 418 IFVSGVQ-AGSPADGQGIQEGDQILQVNDVPFQNLTRREAVQFLGLPPEGEMELVYQK 476
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1251 -----GDVSVVTGHHO--EPASSSLFTGLTSTSIQO--DDLGP 1285
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 477 QDIFMKWQSVRGDSFYIRTHFELEP--SPPSGLGFTRGDVFHVLDTLHP 524
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: July 12, 2001, 14:43:17
Job time: 210 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2001, 14:39:27 ; Search time 64.35 Seconds
(without alignments) 2822.918 Million cell updates/sec

Title: US-09-502-698-1
Perfect score: 6990

Sequence: 1 MWCRRTPPTQSELDLDLDCIELEKPHVDLGEFISSETEDEVLAMTAGOSTEEV 1373

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: SP:archaea:*
2: SP:bacteria:*
3: SP:fungi:*
4: SP:human:*
5: SP:invertebrate:*
6: SP:mammal:*
7: SP:mhc:*
8: SP:organelle:*
9: SP-phage:*
10: SP-plant:*
11: SP-rodent:*
12: SP-unclassified:*
13: SP:vertebrate:*
14: SP:virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6735	96.4	2042	4 075970	075970 homo sapien
2	5811	83.1	2054	11 055164	055164 rattus norv
3	5617	80.4	2055	11 0921K3	0921K3 mus musculu
4	2297.5	32.9	526	11 008783	008783 mus musculu
5	2252	32.2	453	4 043798	043798 homo sapien
6	1820.5	26.0	1134	4 09H3N9	09H3N9 homo sapien
7	1550.5	22.2	1552	4 060833	060833 homo sapien
8	1525.5	21.8	1582	4 043742	043742 homo sapien
9	1463.5	21.0	612	11 070471	070471 mus musculu
10	1390.5	19.9	1524	4 015249	015249 homo sapien
11	1044.5	14.9	2208	5 009515	009515 caenorhabdi
12	589	8.4	728	11 070263	070263 mus musculu
13	564.5	8.1	665	5 023823	023823 calliphora
14	556	8.0	674	5 028BV3	028BV3 drosophila
15	554	7.9	674	5 024008	024008 drosophila
16	534	7.6	1012	5 021075	021075 caenorhabdi
17	517.5	7.4	1551	4 014160	014160 homo sapien
18	507.5	7.3	2460	11 064512	064512 mus musculu
19	507	7.3	2484	6 028006	028006 bos taurus

20	453.5	6.5	1256	4 075085	075085 homo sapien
21	443	6.3	2766	11 09QZB8	09QZB8 rattus norv
22	441	6.3	1150	4 09HBC4	09HBC4 homo sapien
23	441	6.3	1756	5 09VBE4	09VBE4 drosophila
24	432	6.2	1125	4 09H2V6	09H2V6 homo sapien
25	431.5	6.2	1179	11 09JK71	09JK71 rattus norv
26	429.5	6.1	1126	11 09EOJ9	09EOJ9 mus musculu
27	426	6.1	1455	4 060510	060510 homo sapien
28	425	6.1	874	4 09HCD8	09HCD8 homo sapien
29	424	6.1	871	5 09XZ35	09XZ35 drosophila
30	423.5	6.1	1278	4 060434	060434 homo sapien
31	422	6.0	871	5 09NB04	09NB04 drosophila
32	409	5.9	1112	11 09WV01	09WV01 mus musculu
33	399.5	5.7	1131	5 018165	018165 caenorhabdi
34	398.5	5.7	1277	11 088382	088382 rattus norv
35	397.5	5.7	1114	11 09RZ71	09RZ71 rattus norv
36	385	5.5	960	5 09VY25	09VY25 drosophila
37	385	5.5	1171	11 054893	054893 mus musculu
38	381	5.5	950	11 09WU36	09WU36 rattus norv
39	379.5	5.4	1043	11 09WTW1	09WTW1 rattus norv
40	377.5	5.4	822	11 088961	088961 rattus norv
41	376	5.4	1322	11 09QZP6	09QZP6 mus musculu
42	370	5.3	1027	4 09H1X7	09H1X7 homo sapien
43	368	5.3	927	11 062402	062402 mus musculu
44	357	5.1	1112	11 097879	097879 rattus norv
45	353	5.1	1464	5 096782	096782 drosophila

ALIGNMENTS

RESULT 1
ID 075970 PRELIMINARY: PRT: 2042 AA.
AC 075970;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DE MULTI PDZ DOMAIN PROTEIN MUPP1.
GN MUPP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI-Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Eng L., Kravtsovsky G., Clapham P.E.;
RT "Human homolog of MUPP1 protein."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF093419; AAC61870.1; -.
DR HSP: Q12959; 1PDR.
DR InterPro: IPR001478; -.
DR Pfam: PF00595; PDZ: 13.
DR SMART: SM00228; PDZ: 1.
SQ SEQUENCE 2042 AA; 218595 MW; F0E1DD029749A762 CRC64;

Query Match 96.4%; Score 6735; DB 4; Length 2042;
Best local similarity 92.6%; Pred. No. 0;
Matches 1336; Conservative 6; Mismatches 3; Indels 98; Gaps 2;
QY 1 MWCRRTPPTQSELDLDLDCIELEKPHVDLGEFISSETEDEVLAMTAGOSTEEV 60
DB 628 MWCRRTPPTQSELDLDLDCIELEKPHVDLGEFISSETEDEVLAMTAGOSTEEV 687
QY 61 QAPLAMEAGIGHIELEKSGKSGFSLDYOPIDPASVVIIRSLVPGIGIAEKOSRLP 120
DB 688 QAPLAMEAGIGHIELEKSGKSGFSLDYOPIDPASVVIIRSLVPGIGIAEKOSRLP 747
QY 121 GRLEMFVNDVNLNSSLLEAEVALKCAPSGTATGAKPLPLSPREGYSAKEDSLTYP 180
DB 748 GRLEMFVNDVNLNSSLLEAEVALKCAPSGTATGAKPLPLSPREGYSAKEDSLTYP 807

```

QY 181 HSCBAGLADPLRADLALVTNDADLVDESTESPYPSPENDSIYTOASILSLHGSSC 240
    |||||
Db 808 HSCBAGLADPLRADLALVTNDADLVDESTESPYPSPENDSIYTOASILSLHGSSC 867
QY 241 GDGLNYSGLSPSPKDYIENSCLPVDLHMSLELYTONLLERODENTPSVDISMGPAS 300
    |||||
Db 868 GDGLNYSGLSPSPKDYIENSCLPVDLHMSLELYTONLLERODENTPSVDISMGPAS 927
QY 301 GFTINDYTPANAIOOYECENTIVWTESHLPSEYISSAELPSVLPDSAGKSEHLLEBSS 360
    |||||
Db 928 GFTINDYTPANAIOOYECENTIVWTESHLPSEYISSAELPSVLPDSAGKSEHLLEBSS 987
QY 361 LACNAECYMLONVSKESERTINIAKGNSSIGMTVSNANKDGLMIVSIIIGCAISRDR 420
    |||||
Db 988 LACNAECYMLONVSKESERTINIAKGNSSIGMTVSNANKDGLMIVSIIIGCAISRDR 1047
QY 421 IAIIDCJLISNEESTISVTNAQARAMLRRHSLIGPDIKIIVPAHLEEFKISLGQSGR 480
    |||||
Db 1048 IAIIDCJLISNEESTISVTNAQARAMLRRHSLIGPDIKIIVPAHLEEFKISLGQSGR 1107
QY 481 VMAIDFSSYGRDIPLEPEREBSGESELONTAYSMNQPRAVELMREPSKSLGISTV 540
    |||||
Db 1108 VMAIDFSSYGRDIPLEPEREBSGESELONTAYSMNQPRAVELMREPSKSLGISTV 1167
QY 541 GGRMGSRSLNGEVMRGIFIKHVLSDSPAGKNGTLKPDRIVE----- 583
    |||||
Db 1168 GGRMGSRSLNGEVMRGIFIKHVLSDSPAGKNGTLKPDRIVEYDGMRLDASHGQAVEA 1227
QY 584 -----APSGSES 590
    |||||
Db 1228 IRKAGNPVFNVOGIIINRPRKSPLESLHNLVPRYNFSSTNPFADSLQINADKAPSOSES 1287
    |||||
QY 591 EPEKAPLCSVPPPPSAFAEMGSDHTOSASKISQDYKEDDEFGYSMKNIERKGTLTGE 650
    |||||
Db 1288 EPEKAPLCSVPPPPSAFAEMGSDHTOSASKISQDYKEDDEFGYSMKNIERKGTLTGE 1347
QY 651 LHMTELEKSHGSLSLAGNKRSMVSFYIGIDPNGAAGDGRLOLADLELEINGOITLY 710
    |||||
Db 1348 LHMTELEKSHGSLSLAGNKRSMVSFYIGIDPNGAAGDGRLOLADLELEINGOITLY 1407
QY 711 GRSQONASSIICAPSKYKIIIFIRNKDAVNOMAVCPGNAVEPLPSNSNINONKTEPEPVT 770
    |||||
Db 1408 GRSQONASSIICAPSKYKIIIFIRNKDAVNOMAVCPGNAVEPLPSNSNINONKTEPEPVT 1467
QY 771 TSDAAYDLSEFKNYOHLPEKOGGLGIAISEEDTLGVIISKSLTEHGVATDRLKYGD 830
    |||||
Db 1468 TSDAAYDLSEFKNYOHLPEKOGGLGIAISEEDTLGVIISKSLTEHGVATDRLKYGD 1527
QY 831 QILAVDEIVGVPIEFKISLTKAKMTVKLTIAENPDQAVPSAAGASGEKKNSQS 890
    |||||
Db 1528 QILAVDEIVGVPIEFKISLTKAKMTVKLTIAENPDQAVPSAAGASGEKKNSQS 1587
QY 891 LMAVOSGSPREESTIRNRSSTPAIFASDPATCIIIPGCEETIISGRGIGLSIYGS 950
    |||||
Db 1588 LMAVOSGSPREESTIRNRSSTPAIFASDPATCIIIPGCEETIISGRGIGLSIYGS 1647
QY 951 DTLGAFIIEHVEYEGAAKCDGRMAGDQILEVNGIDLKRAKTHDEAIVLRTQORVRLT 1010
    |||||
Db 1648 DTLGAFIIEHVEYEGAAKCDGRMAGDQILEVNGIDLKRAKTHDEAIVLRTQORVRLT 1707
QY 1011 LYRDEAPYKEEVECDTLTIELQKRPKGGLGSIYVGRNDYGVSDIYKGGIADPDGRLI 1070
    |||||
Db 1708 LYRDEAPYKEEVECDTLTIELQKRPKGGLGSIYVGRNDYGVSDIYKGGIADPDGRLI 1767
QY 1071 QGDDIILVNGDVRNAAQOENAAALLKCSIGVTLEVGRIRKAGPHERPSQTSQVSEGS 1130
    |||||
Db 1768 QGDDIILVNGDVRNAAQOENAAALLKCSIGVTLEVGRIRKAGPHERPSQTSQVSEGS 1808
QY 1131 LASFTEPLSGSSTSESSSKKNALASEIOGLRTVEKKKGPDTSLGISTAGVGSPLGD 1190
    |||||
Db 1809 -----GSSTSESSSKKNALASEIOGLRTVEKKKGPDTSLGISTAGVGSPLGD 1859

```

```

QY 1191 VPFIAMMHPTGVAAGTOKLRVGDRIYICSTEGMTHTQAVMLNKASSGSIEMQVAG 1250
    |||||
Db 1860 VPFIAMMHPTGVAAGTOKLRVGDRIYICSTEGMTHTQAVMLNKASSGSIEMQVAG 1919
QY 1251 GDVSVTGHHOEPASSSLSTFGLTSTIFODDLAPPOCKSTTLERGPDLGFSIVGYGS 1310
    |||||
Db 1920 GDVSVTGHHOEPASSSLSTFGLTSTIFODDLAPPOCKSTTLERGPDLGFSIVGYGS 1979
QY 1311 PHGDLPIYKTVFAKGAASEGRLKRGDQIIAVNGOSLEGVTHEAVAILKRTKGTVTLM 1370
    |||||
Db 1980 PHGDLPIYKTVFAKGAASEGRLKRGDQIIAVNGOSLEGVTHEAVAILKRTKGTVTLM 2039
QY 1371 VLS 1373
    |||
Db 2040 VLS 2042

RESULT 2
ID 05164 PRELIMINARY; PRT; 2054 AA.
AC 05164;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE MULTI PDZ DOMAIN PROTEIN 1.
GN MUPPL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathu; Muridae; Murinae; Rattus.
OX NCBI_TaxID-10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE-9819685; Pubmed-9537516;
RA Ulmer C., Schmuck K., Flage A., Lubbert H.;
RT "Cloning and characterization of MUPPL, a novel PDZ domain protein.";
RL FEBS Lett. 424:63-68(1998).
DR EMBL; AJ001320; CA04681.1; -.
DR HSSP; Q12959; 1PDR.
DR InterPro; IPR001478; -.
DR Pfam; PF00595; PDZ; 13.
DR SMART; SM00228; PDZ; 1.
SQ SEQUENCE 2054 AA; 218590 MW; 44BD3F428801F78F CRC64;

Query Match 83.1%; Score 5811; DB 11; Length 2054;
Best Local Similarity 79.5%; Pred. No. 0;
Matches 1147; Conservative 102; Mismatches 116; Indels 78; Gaps 5;

QY 1 MVCCRTVPPTTQSELDLDCDIELEKPHVDLGEFISSETEBDPYLAMTDAGQSTEEV 60
    |||||
Db 620 MVCCRTVPPTTQSELDLDCDIELEKPHVDLGEFISSETEBDPYLAMTDAGQSTEEV 679
QY 61 QAPLAMEAGIOHLELEKSGKGLFSLVDQDPIDPASTVIYIISLYPGGAEDGRLP 120
    |||||
Db 680 QAPLAMEAGIOHLELEKSGKGLFSLVDQDPIDPASTVIYIISLYPGGAEDGRLP 739
QY 121 GDRLMFNDVLENSSLAEAVALKAPSGTVRIGVAKPLPLSPSEGYVSAKEDSFYYP 180
    |||||
Db 740 GDRLMFNDVLENSSLAEAVALKAPSGTVRIGVAKPLPLSPSEGYVSAKEDSFYYP 799
QY 181 HSCBAGLADPLRADLALVTNDADLVDESTESPYPSPENDSIYTOASILSLHGSSC 240
    |||||
Db 800 HSCBAGLADPLRADLALVTNDADLVDESTESPYPSPENDSIYTOASILSLHGSSC 859
QY 241 GDGLNYSGLSPSPKDYIENSCLPVDLHMSLELYTONLLERODENTPSVDISMGPAS 300
    |||||
Db 860 GDGLNYSGLSPSPKDYIENSCLPVDLHMSLELYTONLLERODENTPSVDISMGPAS 918
QY 301 GFTINDYTPANAIOOYECENTIVWTESHLPSEYISSAELPSVLPDSAGKSEHLLEBSS 360
    |||||
Db 919 GFTINDYTPANAIOOYECENTIVWTESHLPSEYISSAELPSVLPDSAGKSEHLLEBSS 974

```

Qy	361	LACNAECWJLWONSKEEFETIINAGNSLIGTWYSAKNDGJMGJIVRSIINGCALSRCR	420
Db	975	LVSABESVTJQNSQOEFEFTVTIANGSSLSIGTWSANKDGLGVIRSTINGCALSRCR	1034
Qy	421	IAIGDCILJINSEESTJISVTNAQARAMLRRISLJPGDIKITYYPARHLEEFKISLQOQGR	480
Db	1035	IAGVDCILJINSEESTJISLTNAQARAMLRRISLJPGDIKITYYPARHLEEFRRSPQOQAG	1094
Qy	481	VMAJDFSSYTGRIPELPEREGEBESELONTAYSNMNPORYELMRPESLGISIV	540
Db	1095	IMAJDFSSYTGRIPELPEREGEBESELONAAVSSWSPRRYELMRPESLGISIV	1154
Qy	541	GGRRMGSRJLNGEVMRCIFIKHJLEDSPPAGKNTLKGRIYE-----	583
Db	1155	GGRRMGSRJLNGEVMRCIFIKHJLEDSPPAGKNTLKGRIYEVGMDLROASHQOAVEA	1214
Qy	584	-----APSOSES	590
Db	1215	IRKAGSVPEVMWOSIVVRPKPSLPJLPHSLYKCFSSSTNPAESLQJLTSKAPSOSES	1274
Qy	591	EPEKAPLCVPPPPSAFAEMGSDHTQSSASKISODVDKDEDEFGYSWKNIERRYGTJGE	650
Db	1275	ESEKATLCVPSPPSVFSEMSDDYQPSATYVAEDEDKEDBEFGYSWKNIORYGTJLQG	1334
Qy	651	LHMLEKSHSGJGLSLAGKKDRSRKSVFTVIGIDPNGAAGKGRJQJADELLEINGQJLY	710
Db	1335	LHMLEKSHSGJGLSLAGKKDRSRKSVFTVIGIDPNGAAGRGRJQJADELLEINGQJLY	1394
Qy	711	GRSHQNASSTIICKAPSKPVKTIIFIRNDAYANOMAVCGNAVEJLPNSSENTLOKKEPEVYT	770
Db	1395	GRSHQNASSTIICKAPSKPVKTIIFIRNDAYANOMAVCGSADJLPSTSEPOKKEPEVPSIT	1454
Qy	771	TSDAANDLSSEKKNVOHLELPRKDDGGJGJIAISEEDTJSGYIJSLETHEGVAATDGLKRGD	830
Db	1455	TSASAVDLSLJNYHLELPRKDDGGJGJIAICEEDJLNGYIJSLETHEGGAADGLKRGD	1514
Qy	831	QJLAVDEIYVGVPIETEFISLTLTAKMTYKLTJHAENPDQAVPSAACAASEKKNSSQS	890
Db	1515	RJLAVDEIYVGVPIETEFISLTLTAKMTYKLTJGAENPDQAVPSAAYTASEKKNSSQS	1574
Qy	891	LMVQSGSPESPERISNRSSRSPATRASPDAPCPIIPGECTTIEISKGTGJLGSIVGS	950
Db	1575	PAPV---APDLEPIPSISKSSTPAIRASDPATCPIIPGECTTIEISKGTGJLGSIVGS	1631
Qy	951	DTLLGAETIHEVYEEGAACKDGRJLMAQDJQJLEVNGIDJLKKATKHADEAINVLRQTPORVJLT	1010
Db	1632	DTLLGAETIHEVYEEGAACKDGRJLMAQDJQJLEVNGIDJLKKATKHADEAINVLRQTPORVJLT	1691
Qy	1011	LYRDEAPYKEEBCVDJLTJELQKKPKGJGLSVGKRNDGYFVSDJYKGGIADPDGRLT	1070
Db	1692	LYRDEAPYKEEBCVDJLTJELQKKPKGJGLSVGKRNDGYFVSDJYKGGIADPDGRLM	1751
Qy	1071	QGGDILLVNGEDVMAQOEVAALLKCSJGTVJLEVGRJAKAPPFISERPSRPSOTSOVSGS	1130
Db	1752	QGGDILLVNGEDVMAQOEVAALLKCSJGTVJLEVGRJAKAPPFISERPSRPSOTSOVSGS	1811
Qy	1131	LJSEFTPLSGSSSTSESSLESSSKNALASEJQJLRTYEMAKGPTSDJGISIAGVSPJGD	1190
Db	1812	LJSEFTPLSGSGSSSTSESSLESSSKNALASEJQJLRTYEMAKGPTSDJGISIAGVSPJGD	1871
Qy	1191	VPIETAMNHPVGAOQJQJRVGDRIYVJCGSTIEBMTHTQAVNLKKNASGSTIEQVYVAG	1250
Db	1872	VPIETAMNHPVGAOQJQJRVGDRIYVJCGSTIEBMTHTQAVNLKKNASGSTIEQVYVAG	1931
Qy	1251	GDVSWVGMHNOERSSSLSTGJLTSISIPDDJLGRPOQSTJLENGRPSLGSIVGGVS	1310
Db	1932	GDVSWVGMHNOERSSSLSTGJLTSISIPDDJLGRPOQSTJLDRPSLGSIVGGVS	1991
Qy	1311	PHGDLPIYVTVFKAASAEDGRLKRGDOJIANQOSLEGVYTHEEVAJLKRKTGTJVL	1370
Db	1992	PHGDLPIYVTVFKAASAEDGRLKRGDOJIANQOSLEGVYTHEEVAJLKRKTGTJVL	2051
Qy	1371	VLS	1373

```

Db      2052 VLS 2054

RESULT      3
09Z1K3
AC      09Z1K3      PRELIMINARY;      PRT: 2055 AA.
DT      01-MAY-1999 (TREMBLrel, 10, Created)
DT      01-MAY-1999 (TREMBLrel, 10, last sequence update)
DT      01-MAR-2001 (TREMBLrel, 16, last annotation update)
DE      MULTIPLE PDZ DOMAIN PROTEIN.
OS      MPDZ.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX      [1]
NCBI_Taxid=10090;

RP      SEQUENCE FROM N.A.
RC      STRAIN=C57/BL6 X CBA F1; TISSUE=WHOLE BRAIN;
RA      Simpson E.H., Suffolk R., Jackson I.J.;
RT      Identification and mapping of mouse Multiple PDZ domain protein,
RT      Mpdz ";
RL      Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AJ131869; CAAL0523.1; -.
DR      HSSP; Q12959; 1PDR.
DR      MGD; MGI:1343489; Mpdz.
DR      InterPro: IPR001478; -.
DR      Pfam: PF00595; PDZ; 13.
DR      SMART: SM00228; PDZ; 1.
SQ      SEQUENCE 2055 AA; 218968 MW; E1A38EE634CC20E4 CRC64;

```

Query Match	80.4%	Score 5617	DB 11	Length 2055
Best Local Similarity	77.1%	Pred No. 0		
Matches 1112	Conservative 114	Mismatches 139	Indels 78	Gaps 5
QY	1	MVCCRRVPTTQSELDLDCDIELTEKPHVDLGEFGSSSETEDEPVLAMTDAGOSTEEV	60	
DB	621	MYCCRRVPTPIALSEMDSLDINDLETERPHLDGFGISSSETEDEPMLAMSDVDQAEEI	680	
QY	61	QAPLAMEGIOHIELEKSKSGLGSIIDYODPIDPASTYIIIRSLVPGGIAEKDORLLP	120	
DB	681	QPLAMEWEGGOSIELEKSKRGLGSIIDYODPIDPANTVIYIRSLVPGGIAEKEDRLLP	740	
QY	121	GRRLFVNDVNLNENSLGEAVELKCAPSGCTRIGAKPLPSPEGGYSAKEDSLTYP	180	
DB	741	GRRLFVNDVNLNENSLGEAVELKCAPSGMVRIGAKPLPSPEGGYSAKEDALCSB	800	
QY	181	HSCEERAGLADKPLFRADLALVGTNDADLVDDESTFESPYSPENDSIYSTOASILSLGSSC	240	
DB	801	HACKESGSLDKALFRADLALIDTPRDESIAESRFESQSPDNDISYSTOASIRSLHDGIC	860	
QY	241	GGGLGGLSSLPSSPPKDYIENSCDPYLDLHMSLEELYTONLERODENTPVSUDISNGPAS	300	
DB	861	SGGMNNGPGLSSPPKDYVSSS-EVYLGLHLSLEELYTONLORHAGSSSDMKPTPTF	919	
QY	301	GFTINDYPTANKEIOOYECCENTIVMTESHLPSEVISAELPSVLPDSAGKSGSEHLEQSS	360	
DB	920	GFXISDYTTNNGEOKYQANPYAMPHSDLPSS-LSFEELASL---AAVAPKYLIDQTF	975	
QY	361	IACNAECVALLQWYSKESPERTINIAKNGSLGATVYSANKDGLMIVRSIIGHGASIRDR	420	
DB	976	LASDAESVYLOGMSQGAEPFRPYTIANKGSSSLGATVYSANKDGLVYIRSIIGHGASIRDR	1035	
QY	421	IAIGDCILSINEESTISVTNAOARMLRHSILGPDIKITTYPAEHLLEEFKISLGOOSGR	480	
DB	1036	IAYGDCILSINEESTISLTNAOARMLRHSILGPDIKITTYPAEHLLEEFRRVFCGOAGG	1095	
QY	481	VALDIFESSYTORDIPELPEREGGESEELONTAAVSNMNOPRVRLMEEPSKISGISIV	540	
DB	1096	IVALDIFESSYTORDIPELPEREGGESEELONTAAVSNMNOPRVRLMEEPSKISGISIV	1155	

```

QY 541 GGRMGSRSLNGEVMRGIFIKHVLKEDSPAGKNGTLKPGDRIVE----- 583
DB 1156 GGRMGSRSLNGEVMRGIFIKHVLKEDSPAGKNGTLKPGDRIVEVGMGLRDSHQAWEA 1215
QY 584 -----APSGSES 590
DB 1216 IRKAGNPFVFNQSIINRPRKSPPLSLPHSLYPKSYFSSSTNPFADSLQTLTDOAPSOSES 1275
QY 591 EPEKAPLCVPPPPSAFAEKSDHTOSASKISODYKDEDEFGYSMKINRERGTTLTGE 650
DB 1276 ETERPALCNVPPSPSVSEKSGDCAQSAFAVEDEKEDFEYSKKNIOERGSGLTQ 1335
QY 651 LHMTELEKSHGSLGSLAGNKRDSRMVFIYIDPNCAGKDGRLQIADLEILEINGQILY 710
DB 1336 LHVTELEKSGSLGSLAGNKRDSRMVFIYIDPNCAGKDGRLQIADLEILEINGQILY 1395
QY 711 GRSNONASSIICAPSKVKIIFIRNKDQVNMAYCPGNAVPELPSNSNINONKTEPEVT 770
DB 1396 GRSNONASSIICAPSKVKIIFIRNKDQVNMAYCPGNAVPELPSNSNINONKTEPEVT 1455
QY 771 TSDAAYDLSPFKYQHLPEKDGGLGIAISEEDTSLGVIKISLTHGVAATDRLKVG 830
DB 1456 TSASADLSSLTDYVQLELPDQGLGIAICEEDTINVMIESLTHGCAKDGRLKPGD 1515
QY 831 QILAVIDEIVGYPIEFKISLTKAKTKVLTIAENPDQAVPSAGASGEKKNSQS 890
DB 1516 HILAVDEVVAGCEVEKISLTKAKTKVLTIAENPDQAVPSAGASGEKKNSQS 1575
QY 891 LMPVOSGPEESIRNRSSTPAIFASDPATCPIPECETTIEISKGRGLGSIYGS 950
DB 1576 PAVP---APDLEPIPSRSRSTPAVPSADPATCPIPECETTIEISKGRGLGSIYGS 1632
QY 951 DTLGAFIHEVEBGAACKDGRIMAGDQILEVNGIDLKATHDEAINVLRQTPORVRLT 1010
DB 1632 DTLGAFIHEVEBGAACKDGRIMAGDQILEVNGIDLKATHDEAINVLRQTPORVRLT 1692
QY 1011 LYRDEAPYKEEVEDTLTIELQKRPKGLGSIYGRNDTGVFVSDIYKGIADPDGRIL 1070
DB 1693 LYRDEAPYKEEVEDTLTIELQKRPKGLGSIYGRNDTGVFVSDIYKGIADPDGRIL 1752
QY 1071 QGDDIILVNGEDVNNASOEAVALLKCSAGTVLEVGRIKAPFHSERPSOTSQVSSGS 1130
DB 1753 QGDDIILVNGEDVNNASOEAVALLKCSAGTVLEVGRIKAPFHSERPSOTSQVSSGS 1812
QY 1131 LSSTFPLSGSSTSESSSKNALASEIOTGLRTEVEMKKGPTDSLIGISAGVSPGLD 1190
DB 1813 LSSTFPLSGSSTSESSSKNALASEIOTGLRTEVEMKKGPTDSLIGISAGVSPGLD 1872
QY 1191 VPIRIAMHPTGVAAGTOQKLRVGRITVTCSTEGMTHQAVNLKKNASGSIEMOVYAG 1250
DB 1873 VPIRIAMHPTGVAAGTOQKLRVGRITVTCSTEGMTHQAVNLKKNASGSIEMOVYAG 1932
QY 1251 GDVSVVGHQOELNAPCLAFGLTSSSTFPDDLSPQSQTITLDRGPDGSLFNVGIGS 1310
DB 1933 GDVSVVGHQOELNAPCLAFGLTSSSTFPDDLSPQSQTITLDRGPDGSLFNVGIGS 1992
QY 1311 PHGDLPIYKTVFAKGAASEDRKLRGDOIIVANGOSLEGVTHEEVAVALIKRTKGTVTLM 1370
DB 1993 PHGDLPIYKTVFAKGAASEDRKLRGDOIIVANGOSLEGVTHEEVAVALIKRTKGTVTLM 2052
QY 1371 VLS 1373
DB 2053 VLS 2055

```

```

GN 99P-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Lee S.S., Weiss R.S., Javier R.T.;
RL Submitted (Apr-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF000168; AA57835.1; -.
DR HSSP; Q12939; 1PDR.
DR InterPro; IPR001478; -.
DR Pfam; PF00595; PDZ; 4.
DR SMART; SM00228; PDZ; 1.
FT NON_TER 1
SQ SEQUENCE 526 AA; 54948 MW; 3843C5E3A2A10E8 CRC64;

Query Match 32.9%; Score 2297.5; DB 11; Length 526;
Best Local Similarity 86.6%; Pred. No. 2.8e-129;
Matches 457; Conservative 34; Mismatches 32; Indels 5; Gaps 2;

QY 848 FISLTKAKTKVLTIAENPDQAVPSAGASGEKKNSQSILMPVOSGPEESIRNT 907
DB 2 FISLTKAKTKVLTIAENPDQAVPSAGASGEKKNSQSILMPVOSGPEESIRNT 958
QY 908 SRSSTPAIFASDPATCPIPECETTIEISKGRGLGSIYGSPTLLGAFIHEVEBGA 967
DB 59 SRSSTPAIFASDPATCPIPECETTIEISKGRGLGSIYGSPTLLGAFIHEVEBGA 118
QY 968 ACKDGRIMAGDQILEVNGIDLKATHDEAINVLRQTPORVRLTLYRDEAPYKEEVEDTL 1027
DB 119 ACKDGRIMAGDQILEVNGIDLKATHDEAINVLRQTPORVRLTLYRDEAPYKEEVEDTL 178
QY 1028 TIE--LQKRPKGLGSIYGRNDTGVFVSDIYKGIADPDGRILQDQIILVNGEDVYRN 1085
DB 179 TIEQLQKRPKGLGSIYGRNDTGVFVSDIYKGIADPDGRILQDQIILVNGEDVYRN 238
QY 1086 ASQAVNALKCSAGTVLEVGRIKAPFHSERPSOTSQVSSGSLSSTFPLSGSSTSE 1145
DB 239 ASQAVNALKCSAGTVLEVGRIKAPFHSERPSOTSQVSSGSLSSTFPLSGSSTSE 298
QY 1146 SLESSSKNALASEIOTGLRTEVEMKKGPTDSLIGISAGVSPGLDVPFIAMHPTGVA 1205
DB 299 SLESSSKNALASEIOTGLRTEVEMKKGPTDSLIGISAGVSPGLDVPFIAMHPTGVA 358
QY 1206 QTOKLRVGRITVTCSTEGMTHQAVNLKKNASGSIEMOVYAGGVSVYTGHOELAN 1265
DB 359 QTOKLRVGRITVTCSTEGMTHQAVNLKKNASGSIEMOVYAGGVSVYTGHOELAN 418
QY 1266 SLSFTGLTSSSTFPDDLSPQSQTITLDRGPDGSLFNVGIGSPTLLGAFIHEVEBGA 1325
DB 419 PCLAFGLTSSSTFPDDLSPQSQTITLDRGPDGSLFNVGIGSPTLLGAFIHEVEBGA 478
QY 1326 GAASEDRKLRGDOIIVANGOSLEGVTHEEVAVALIKRTKGTVTLMVLS 1373
DB 479 GAASEDRKLRGDOIIVANGOSLEGVTHEEVAVALIKRTKGTVTLMVLS 526

RESULT 5
O43798 PRELIMINARY; PRT; 453 AA.
AC O43798;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE MULTI PDZ DOMAIN PROTEIN 1 (FRAGMENT).
GN MUPP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;

```



```

Db 863 -YHSTDADFTGCGFOAPLAV-----DPATCPVIGQEMITIEISGRSGIG 907
OY 944 LSTVGGSDTLTGAFIIEHYVEEGAACKDGRUMAGDOLLENGIDLRKATHEALNVLROT 1003
Db 908 LSTVGGSDTLTGAFIIEHYVEEGAACKDGRUMAGDOLLENGIDLRKATHEALNVLROT 967
OY 1004 PORVRLTLRYDEAPYKEEVCDTLTTELQKKPKGKGLSIVGRKNDGVFVSDIVKGGIA 1063
Db 968 POKRVLVYRDEAHYRDEENLEIFPVDLQKKAGKGLSIVGRKNGSGVFIISDLVKGGA 1027
OY 1064 DPGRLLIQQDQILLYNGEDVBNASQEAVALKCSLGTVTLVGRKAGPFRSPSQT 1123
Db 1028 DLDRLIQQDQILLYNGEDVBNASQEAVALKCSLGTVTLVGRKAGPFRSPSQT 1087
OY 1124 SQVSE 1128
Db 1088 SQVSE 1092

RESULT 7
OY 060833 PRELIMINARY; PRT; 1552 AA.
AC 060833;
DT 01-AUG-1998 (TEMBLrel. 07, Created)
DT 01-AUG-1998 (TEMBLrel. 07, Last sequence update)
DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
DE INADL, C-TERM VARIANT2.
GN INADL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homindae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Philipp S.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=97424368; PubMed=9280290;
RX PHILIPP S., Flockerzi V.;
RT "Molecular characterization of a novel human PDZ domain protein with
RT homology to INAD from Drosophila melanogaster.";
RL FEBS Lett. 413:243-248(1997).
DR EMBL; AJ224747; CAA12112.1; -.
DR HSSP; P31016; 1BE9.
DR InterPro; IPR001478; -.
DR Pfam; PF00595; PDZ; 8.
DR SMART; SM00228; PDZ; 1.
DR SRA; SRA00228; PDZ; 1.
SQ SEQUENCE 1552 AA; 170322 MW; FF80D9A65350F996 CRC64;

Query Match 22.2%; Score 1550.5; DB 4; Length 1552;
Best Local Similarity 37.3%; Pred. No. 8e-84;
Matches 419; Conservative 131; Mismatches 293; Indels 281; Gaps 36;
OY 1 MWCRRTPPTTOSLSDLCITELTEKPHVDGEFISSETPVLAAMDAGOSTEEV 60
Db 633 LVCCRRLE--DDEASVDEPRTERSTLPE--TEVDHMDVNTDEEDD----- 674
OY 61 QAPLAMEAGIQTIELKSGKGLGFSILDYQDPDPASTYIIIRSLVPGIAGKGRLLP 120
Db 675 -GELALMSPEKIVELVKDKGLGFSILDYQDPDPASTYIIIRSLVPGIAGKGRLLP 733
OY 121 GDRLMFVNDVLENSLEAVEALKGAPSGVIRGVAKPLPLSPGEG-----YVSAKED- 174
Db 734 GDRVAVSNEKRLDNTSLAEVAEILKAVPGLVHGLCKPLVEDNEEESCITLHSSNEDK 793
OY 175 -SFLYPPHSCSEA-----GLADKPLRADLALVGTNDALVDLVESTFESPSPENDSIY 226
Db 794 TFEFGTTHDINSILILAPAGFRODEPYFK-----BELVDE-----PFLDLGKSPH 838
OY 227 STQASITLSLGGSGDGLNGSSLPSPKDVLENSCDPVLDMHSLLEYLTONTLEROD 286

```

```

Db 839 SQKEI-----RQSKAEV-----KHEFLTRLOQWDE 866
OY 287 E-----NTPSVDISMGPAAGFTINDYIPANAEIOYCEMTIWTESHLESEV 334
Db 867 EREMLVDEEYELVODPSPSMELPLS--HQEATPVSVN--ELHFGTOMLHNEPSES 921
OY 335 ISSAELPSVLPDSAGKSEHLLQEOSLACNAECVMLONVSKESPRTINIAKNGSSGMT 394
Db 922 QEATGRTVYSGEA-----QPYGVC--PENMKENFWESLP--SVPTSEGSOGR- 969
OY 395 VSANKDGLMIVRSIIHGALSRDGRILAGDCILISINEESTISVTNAQARMLRHSILG 454
Db 970 -FDLEMLNLSIAKTSILDGMIPND-----VQG 995
OY 455 PDIKITY-VPAHELEEFKISLQ--QSGRVNADLFSSYTG-----RDIPELPEREGE 505
Db 996 PSLILDLPPVAQREQEDLPYHOATRVISK--ASAVTGLMSRYATDTCELPEREGE 1053
OY 506 GESELONTAVSNMOPRVELMREPESKLSISIVGRMGSRLSNCEVNRGIFIKHYLE 565
Db 1054 GEET-----PNFHWGPRRIYEIFREPVSIGISIVGQYVYKRLKNEELAKGIFIKYLE 1109
OY 566 DSPAGKNGTLKPGDRIYEA-----PSQSESEPEKA-----PLCSVPPPPSA 607
Db 1110 DSPAGKTNALKTGDKILEVSGVDLQNSHBAVEAIKKNAGNPVYFIVQSLSFPRVIVN 1169
OY 608 FAENG-----SDHTQSSASK-----ISQDVK--DDEGYSKKNRER 643
Db 1170 HNKANKITGONODTQKKERKROGTAPPMKLPAPPYALYDDSDNEEDAEFTDQKTROR 1229
OY 644 YGTLTGELHMELEKSGHGLSLAGNKRDSRMSVFYIGDIPNGAAGDRLQIADLELLE 703
Db 1230 YADLPGLHIIIELEKDKNGELSLAGNKRDSRMSVFYIGDIPNGAAGDRLQIADLELLE 1289
OY 704 INGOILYGRSHQNASSTIKCAPSKVKIIFIRNKDAVQMAVCPGNAVEPLPSNSENLOK 763
Db 1290 INNOILYGRSHQNASSTIKCAPSKVKIIFIRNKDAVQMAVCPGNAVEPLPSNSENLOK 1340
OY 764 ETEFTVYTSAAVDLSFKNVQHLELPKDDGGLGIALISEDTLSGVLIKSLTEGVNATD 823
Db 1341 -----SST-----EDSGTEPISSSE-----D 1357
OY 824 GRUKVGDQILAVDEIVGVPIEKFISLTKAKTVLTTIHAENPDQAAPSAAGASGE 883
Db 1358 GSELEVGIKQLPSESEFKLA-----VSQMKQOKFTPVSF-----SSQELPLAPASS- 1403
OY 884 KKNSSQSLMVPQSGSPESIRNTSRSTPAIFASDPATCPPIIPGCTTIEISGRGTGLG 943
Db 1404 -YHSTDADFTGCGFOAPLAV-----DPATCPVIGQEMITIEISGRSGIG 1448
OY 944 LSTVGGSDTLTGAFIIEHYVEEGAACKDGRUMAGDOLLENGIDLRKATHEALNVLROT 1003
Db 1449 LSTVGGSDTLTGAFIIEHYVEEGAACKDGRUMAGDOLLENGIDLRKATHEALNVLROT 1508
OY 1004 PORVRLTLRYDEAPYKEEVCDTLTTELQKKPKGKGLSIVGRK 1047
Db 1509 POKRVLVYRDEAHYRDEENLEIFPVDLQKKAGKGLSIVGRK 1552

RESULT 8
OY 043742 PRELIMINARY; PRT; 1582 AA.
AC 043742;
DT 01-JUN-1998 (TEMBLrel. 06, Created)
DT 01-JUN-1998 (TEMBLrel. 06, Last sequence update)
DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
DE INADL, C-TERM VARIANT3.
GN INADL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homindae; Homo.
OX NCBI_TaxID=9606;

```

QY	DB	1290	YADPBGELHIIELEKDKNGLGISLGNKDRMSITFVVGINFGPAAADGRRIIGDELLE	1289
QY	704	INGOILYGRSHONASSIITKCAPSKVYIIFIRNKDAVNOMAACPNAVEPLPSNSENTLONK	763	
DB	1290	INNOILLYGRSHONASAIITKTAAPSKVLYEIRNEDAVNOMAVP----	1340	FPVPSSSP-----
QY	764	EETPVTTSDAAVLSSFRKNVQHLELPKDDGGGLGIAISEDTLGSVYIKSLTEHGVATD	823	
DB	1341	-----SST-----EDOSTEPISSE-----D	1357	
QY	824	GLKVGDDHLLAVDEIVVGVPIETKFTSILTKTKMIVKLIHANEPSDAVPNAGASGE	883	
DB	1358	GSLEVGITQOLPESSEFKLA-----VSQKKQKYPTRKVS-----SSOEIPLAPSS---	1403	
QY	884	KKNSQSLSLWPDGSGSPESPIRNTSRSSSTPAIFASDPATCPIIPCCETTEIISKGTGIG	943	
DB	1404	-YHSYDAFTGTCGQNALSV-----DPATCPIIPQEMIEIISKSGSGIG	1448	
QY	944	LSIVGSDTL-----GAFIIEVEEGCAACKGR	973	
DB	1449	LSIVGKDTPLFWRLGSPRAMSQHLVRAFMILHPYTEVEGQNAIYIHEVEEGGAARDR	1508	
QY	974	LMAGQIILEVNGIDLRKRPDEAIIVNLQTPORVNLTLTYRDAPYKEEVCPTLTIELQ	1033	
DB	1509	LMAGQIILEVNGVDLRNSSHETALRQTPOKVRLVYRDEAHYRDEENLTFPVDLQ	1568	
QY	1034	KPGKGLGLSIYGR 1047		
DB	1569	KAGRGGLSIYGR 1582		
RESULT	9			
ID	070471	PRELIMINARY:	PRT:	612 AA.
AC	070471			
DT	01-AUG-1998	(TREMblrel. 07, Created)		
DT	01-AUG-1998	(TREMblrel. 07, Last sequence update)		
DT	01-MAR-2001	(TREMblrel. 16, Last annotation update)		
DE	CHANNEL-INTERACTING PDZ DOMAIN PROTEIN.			
GN	CIPP.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BAIB/C; TISSUE=CEREBELLUM, BRAIN;			
RA	Kirschner C., Mermelstein P.G., Holden W.T., Summelfer D.J.;			
RL	Mol. Cell. Neurosci. 0:0-0(1998).			
DR	EMBL; AF060539; AAC40148.1; -			
DR	HSSP; Q12959; 1PDR.			
DR	MGI; MGI:1277960; CIPP.			
DR	InterPro; IPR001478; -			
DR	Pfam; PF00595; PDZ, 4.			
DR	SMART; SM0228; PDZ, 1.			
SO	SEQUENCE 612 AA; 65420 MW; 145DE1769F54BE56 CRC64;			
Query Match	21.0%;	Score 1469.5;	DB 11;	Length 612;
Best local Similarity	49.5%;	Pred. No. 1.2e-79;		
Matches 335;	Conservative 85;	Mismatches 158;	Indels 99;	Gaps 12
QY	638	KNIRREYGTILTELMIIELEKSGISGLSLGNKDRSRMSVYIGIDPAGCAKDRLOI	697	
DB	8	EKIRRYAVDLPELHIIELEKDKNGLGISLGNKDRSRMSITFVVGINFGPAAADGRRI	67	
QY	638	ADELLEINGOILYGRSHONASSIITKCAPSKVYIIFIRNKDAVNOMAACPNAVEPLPSNS	757	
DB	68	GDELEINNOILYGRSHONASAIITKTAAPSKVLYEIRNEDAVNOMAVP-----	116	
QY	758	EYLQKKEPEYVTTSDAANDLSSFRKNVQHLELP-KDQGLGIAISEDTLGSVYIKSLTE	816	
DB	117	-----PELISHSPSPVDLGGTELVSSBE-----SSYDAKHLPE	151	

```

QY 817 HGVAATDGLRLAVGDOILAVDEIVGVPIEKFISILTKAKMTVKLTIHAENPDSOAQVSA 876
D 152 PESS-----KPELDSQVVDNMVAEQKE-----SESPDSACQIK 187
QY 877 AGAAGEKKNSSQSLMVPQSGSPPESTIRNTSRSS-----TPAIFASDPATCPPIIP 927
D 188 QQSTVSSTSSSSOD--SPSSPAPLQSAHADVDSGNFQAPLDPAPLPSVDPATCPPIIP 245
QY 928 GCETTFISKGRTRGLSIVSGSDTLTGAFTHIEVEGGAACKDGRMLAGQOILLEVNGID 987
D 246 GQEMTEIISKGRSLGSIYSGKDTPLDAIVHEVEGGAARGRMLAGQOILEVNGVD 305
QY 988 LRKATHEAIVNLKOTQPVRLTYLRDEAPYKEEVCDTLTELOKPKGKGLSIVGKR 1047
D 306 LRSSSHEAIVNLKOTQPVRLTYLRDEAPYKEEVCDTLTELOKPKGKGLSIVGKR 365
QY 1048 NDIGVAVSDIYKGGIADPDGRLIGDQILLVNGEDVRNASEAVAAALIKCSLGTVLEVG 1107
D 366 SGSGVFISDIYKGGIADPDGRLIGDQILLVNGEDVRNASEAVAAALIKCSLGTVLEIG 425
QY 1108 RIRKGPHESERPSTQSVSGSLSTFPLSGSTSESLSS-----SKKNMLASRIQ 1161
D 426 RLRAGSMAASKTQSONSGDQSHASHSCRP--SEAPVITSLQNLVGTKRSDPQKCTE-E 483
QY 1162 GLRTVEKKKGFDSLIGISIVAGVSPGLADYPIFTAMMHPTGVAQOTKRLARDRIIVTICG 1221
D 484 EPRVEIIRLSDALGISIAGCKGSPGLGDIRIFTAMIQANGVAARTOKLAKGDRIVSING 543
QY 1222 TSTGMRHTQAVNLKNAKSGSIEMQVAVGADVSVTGHQHPASSSLFTGLTSTSFOD 1261
D 544 QPLGLSHTDVAIVNLKNAKSGSIEMQVAVGADVSVTGHQHPASSSLFTGLTSTSFOD 589
QY 1282 DLGPPCKSITLERGPD 1298
D 590 QLGSP-----TADRHPE 601

RESULT 10
015249 ID 015249 PRELIMINARY; PRT; 1524 AA.
AC 015249:
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE PDZ DOMAIN PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97424368; Pubmed=9280290;
RA Philipp S., Flockerzi V.;
RT "Molecular characterization of a novel human PDZ domain protein with
RT homology to INAD from Drosophila melanogaster.";
RL FBS Lett. 413:243-248(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96224170; Pubmed=8617505;
RA Lennon G., Aufiray C., Polymeropoulos M., Soares M.B.;
RT "The I.M.A.G.E. Consortium: an integrated molecular analysis of
RT genomes and their expression.";
RL Genomics 33:151-152(1996).
DR EMBL: AJ001306; CA04666.1;
DR HSSP: P31016; 1BFE
DR InterPro: IPR001478;
DR Pfam: PF00595; PDZ; 8.
DR SMART: SM00228; PDZ; 1.
SO SEQUENCE 1524 AA; 167323 MW; E3BBA7EC856A954D CRC64;

```

Query Match 19.9%; Score 1390.5; DB 4; Length 1524;

```

Best Local Similarity 35.1%; Pred. No. 2,9e-74;
Matches 395; Conservative 126; Mismatches 262; Indels 309; Gaps 37;

QY 1 MYCCRRVPPPTQSELDLDCIDLETEKPYVDLGEFGSSEDEPVLAMTDAGOSTEYV 60
D 633 LVCCRRLP--DDEASVDEPRRTETSLPE--TEVDHMDVNEDEDD----- 674
QY 61 QAPLAMEAGIOHIELEKSGSLGSLDYODPIDPASTVIRISLVGGIAEKDGRLLP 120
D 675 -GELAMSPREKIVELVYDCDCKGLGSLDYODPIDPASTVIRISLVADGAESGGLP 733
QY 121 GDRLMFVNDVNLMSLEAEVALKGAESTGRIGVAKPLPSPEG-----YVSAKED 174
D 734 GDRLVSVNEYRLDNTSLAEAVEILKAAPPGLVHLGICRPLVEDNEESCYLIHSSNEDK 793
QY 175 -SFLYPRSCSEA-----GLADKPLRALDALVGTADALVNESTESYSPENDSIY 226
D 794 TEFSGTHIDINSLILEAPKFRDEPYFK-----EELVDE----PFLDLSNFH 838
QY 227 STQASILSLHSGSGDGLMYGSSLPSPPKDVIESCDPVLDMHLSBELYTONLLEROD 286
D 839 SQCKEI-----EOKRAWE-----MHFEITPLQEMDE 866
QY 287 E-----NTPSYDISMPASGFTINDYPPANAIEQYECENTIVFTESHLDBEV 334
D 867 EREMLVDEEYELDYODPSRSMELYPLS--HIOEATPPVPSV--ELHFGTQWLHNPSES 921
QY 335 ISSALPVLVLDSDSGKSGEHLLEOSLACNAECVLOWNSKESPERTINIKNGSSIGMT 394
D 922 QEARTGRIVSOEA-----OPGYTC--PENWKEFWESLP--SVPTSGNSQOCR- 969
QY 395 VSANKDGLMIVRSIIHGAIISRDRGRIAGDCILSINEESTISVTNAQARMLRHSILIG 454
D 970 -FDDLENLNLSAKTSLDGLMIPND-----VQG 995
QY 455 PDIKITY-VPAHELEEFKISLQ--QSGRYVALDIFSSITG-----RDLELPEREGE 505
D 996 PSLILDLPPVNAQRREOEDLPYQHOATRVISK--ASAYTGLMSRYAVDTCELEPEREGE 1053
QY 506 GEESELQNTVAANNPPRVELMREPSKLSISIVYSGMGKSLNCEVGMGIFIKHYLE 565
D 1054 GEET-----PNSHNGPPRIIVIFREPNNVSLGISIVGGQTVIKRLKNGELKGITIKOYLE 1109
QY 566 DSPAGKNGTLKPGRIYEA-----PSQSESEPEKA-----PLCSVPPPPSA 607
D 1110 DSPAGKNTALKTGDKILEVSGVDLQNASHSBAVAIKNAGNPVYIYQSLSTSPRVIP 1169
QY 608 FAEMG-----SDHTQSSASK-----ISQDVVK--EDEFYSKKNIRER 643
D 1170 HNKANKITGNQNOPTQEKREKRGCTAPPKMLPPPYKALITDSDNEEDAFDQKIRQR 1229
QY 644 YGTLTGLHMLIEKRGHSLGLSLAGNDRSRMSVFIYDIPNAGAKDGLQIADLELLE 703
D 1230 YADLPGLHITELKDKNGDLGLSLAGNDRSRMSIFVYGVINPEGPAAADGMHIDELLE 1269
QY 704 INQILYGRSHQNASIITKCAPSKVYKIIIFRNKDAVNOMAVCPGNAVEPLPSNSENLQNK 763
D 1290 INNQILYGRSHQNASIITKCAPSKVYKIIIFRNKDAVNOMAVCPGNAVEPLPSNSENLQNK 1340
QY 764 ETEFTVITSDAAVLDSSFKNVQHLELPKDDGGLGIAISEPTLSGVIIKSLTEHGVATD 823
D 1341 -----SST-----EDSGTEPISSE-----D 1357
QY 824 GLKLVGDOILAVDEIVGVPIEKFISILTKAKMTVKLTIHAENPDSOAQVSAAGSAGE 883
D 1358 GSLVGVKQLPSESFKLA-----VSOMKQKIPYVVSF-----SSEITPLAASS--- 1403
QY 884 KKNSSQSLMVPQSGSPPESTIRNTSRSTPAIFASDPATCPPIIPGCETTFISKGRGLG 943
D 1404 -YHSTADDFGCGFGQAPLVS-----DPATCIYVGGQEMTIRISKRSLG 1448
QY 944 LSIYGSDDLIGAFTHIEVEGGAACKDGRMLAGDQILLEVNGIDLRKATHEAIVNLKOT 1003

```



```

Db 1977 GKVALKTEENNNDPVRYAVAVASPPAATPSTASVYITSAPIVSPEDNNVVEPIPGTIS 2036
| | | | |
Oy 1115 -----HERRPSQTSOYSEGSLSFFP-----PLS 1139
| | | | |
Db 2037 AREERNDDVPPAPMRPIITHTTSPGECLEFQOEPAGLSPTEPSESSGCGFMVSQVEEOPRT 2096
| | | | |
Oy 1140 GSSTSELESSSKKNALASPI-----OGLRFT--VEMKGPTDSIGSIAGVSGPLGDV 1191
| | | | |
Db 2097 PTTT-----SSNNNNNSLAIDIHDLKEBSDDLVLVELKRVYDQGLM-----GIGK--RSR 2147
| | | | |
Oy 1192 PPIAMMHPGVAAGTOKLAVGDRIVTICGTSTEGCTHTQAVNLKKNAGSIEMOV 1247
| | | | |
Db 2148 GILVTSLOPQSAAA--EKLKVGGRILAV--NALPVSDQLSAVTFVKAGSGORLYLDI 2199
| | | | |
RESULT 12.
O70263 PRELIMINARY: PRT: 728 AA.
AC 070263; 070264;
O1-AUG-1998 (Tremblrel, 07, Created)
O1-AUG-1998 (Tremblrel, 07, Last sequence update)
O1-MAR-2001 (Tremblrel, 16, Last annotation update)
DE NUMB-BINDING PROTEIN LNX (LIGAND OF NUMB-PROTEIN X).
CN LNX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN 11
RS
RC SEQUENCE FROM N.A. AND CHARACTERIZATION.
RP TISSUE-BRAIN, AND EMBRYO.
RX MEDLINE=98204916; PubMed=9535908;
RA Dho S.E., Jacob S., Wolting C.D., French M.B., Rohrschneider L.R.,
RA MGladie C.J.;
RT "The mammalian numb phosphotyrosine-binding domain. Characterization
RT of binding specificity and identification of a novel PDZ domain-
RT containing numb binding protein, LNX."
RT J Biol Chem. 273:9179-9187(1998)
CC -1- FUNCTION: INTERACTS WITH THE PHOSPHOTYROSINE INTERACTION DOMAIN OF
CC THE NUMB PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS; ISOFORM LNX (OR LNXP80) (SHOWN
CC HERE) AND ISOFORM LNX-B (OR LNXP70); ARE PRODUCED BY ALTERNATIVE
CC SPLICING.
CC -1- TISSUE SPECIFICITY: BOTH ISOFORMS ARE EXPRESSED IN THE HEART.
CC ISOFORM LNX IS ALSO EXPRESSED IN KIDNEY, LUNG AND SKELETAL MUSCLE
CC WHILE ISOFORM LNX-B IS ALSO EXPRESSED IN BRAIN.
CC -1- SIMILARITY: CONTAINS 4 PDZ DOMAINS.
DR EMBL: AF034745; AAC40075.1; -
DR EMBL: AF034746; AAC40076.1; -
DR HSSP: P31016; 1BFE.
DR MGD: MGI:1278335; Lnx.
DR Interpro: IPR001478; -.
DR Pfam: PF00097; zf-C3HC4; 1.
DR Pfam: PF00595; PDZ; 4.
DR SMART: SM00228; PDZ; 1.
KW Alternative splicing.
FT DOMAIN 278 362 PDZ.
FT DOMAIN 385 467 PDZ.
FT DOMAIN 508 593 PDZ.
FT DOMAIN 638 723 PDZ.
FT MOTAGEN 188 188 Y->A: ABOLISHES BINDING TO NUMB PROTEIN.
FT MOTAGEN 188 188 Y->F: NO EFFECT ON BINDING TO NUMB
FT VARSPLIC 1 131 PROTEIN.
FT NM0PLADDPPSPPLCTVCGNHSHPENHFTYTEDVDD
FT DLIQICQALLDPLDPGCGTYCTCLDITLNFVLEKDCPD
FT KPELVQCKKSSITLVNKLKLAVTQPFDEHCEVLQRD
FT LQHHQTS -> MKALLLVLPWLSPPANYIDWNGNLFYLS
FT EL (IN ISOFORM LNX-B).
SO SEQUENCE 728 AA; 80156 MW; E2914BD364CCEC4 CRC64;
Query Match 8.4%; Score 569; DB 11; Length 728.

```

Best Local Similarity 29.1%, Pred. No. 6.9e-27; Matches 188; Conservative 115; Mismatches 238; Indels 106; Gaps 23;

```
OY   779 SSKNVQHLLEPKD-----GGGL--GAISEEDLISGYIILSTEHGAATDGRLKVDQ 831
      :|::|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db    130 TSCCKASHYGTLTKRKRRRSODGPCGCASIMLATLS-----PEVSAAR----- 173
OY   832 ILAVDEIVGVPIKEEFTSLKTAKMTVKLIHMEKDPAVP---SAGAASEKKRSS 888
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    174 ISLMTEGCLDNF--AYVSVEDEGPVANSSDSGRSNTRAPFERBMSRKRKTINRA 231
OY   889 QS-LMWPOSGSPDEPESI---RMTRSRTPAIFADPATCPIIDCE-TTIEISKRGTL 942
      ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    232 LSLALERTKSGSVVAHVNDQGRDNSNTPTPEVF---PLFLPIDGETTSIKIRAPSE 288
OY   943 GLSI--VGSSDTLGAFIIHEHYEGCAACKXGRIAMAGOLILEVNGIDLKRAFTHEDAINVL 1000
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    289 SLSTLVGSSESPLVIHIITQHTRGVTAARGRLPGIILLKVGMDSNVPHTAVRL 348
OY   1001 ROTPORVALTYLRDE-----APYKEEVCDLTLETLOK-RPKGGLGISYKERNDT 1050
      ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    349 RQPCCVRLTVLREQKFRSRNAHPDYSYPRDDSFHILNKSSPEEQIGIKLVRPDER 408
OY   1051 GFVFSVDLYKGGLADPDGLIOGDQILVNGEDVRNASGEANALLKCSLTGYTFEYGRIK 1110
      ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    409 GFETINVINGVADRHGOLEDENDRYLLAHGHDLRGPESAAHLIQSERVVHLVSR-- 466
OY   1111 AGPFSEHRPS----QTQVSEGSLSTFPPLSGSSSTSELESSKNMALASEIOGLRT 1165
      :||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    467 -----QVHQSSPDIREQAEGWISNGQS---PGGER-----NMASKPAATCHE----KV 508
OY   1166 VEMKGPFDLSIGISTAGVGPLDVPEFFIAMHPGTVAOTOKLRNGDRIVTCGSTE 1225
      |:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    509 VSVMWDPPESLKMTYVGSHAREMDLPITYIVSEBGVISHDGIKGDILLANGIELT 568
OY   1226 GMTTHQAVNLKNAAGSILEMOVVAAGDVSVYTGHNOEPASSLSFTGILTSTIPDDGLP 1285
      :||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    569 EYSREAEVAILKSAPSYYALKAL---EK-----EQDAQEDCPALADS-----NNHWTP 615
OY   1286 P-----OQKSTITLERGPDG-IQFSYVGGYGSPHGDDLPIYKVTYRAKG 1326
      :||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    616 PGDWSPSWMMLELPQLCYLCNCRDYLRENTPGSLGFICYEEYSGMKPFIKSIGEST 675
OY   1327 AASESGRLKRGDIIVANGOSLEGVTHEEAVALIKRTGTYTLAWLS 1373
      ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    676 PAYNDGRICGDILLAVNGRSTSGIHACLAAMLKELKGRITTIITAS 722
```

RESULT 13
Q23823 PRELIMITARY; PRF; 665 AA.

AC Q23823:
DT 01-NOV-1996 (TREMBREL_01, Created)
DT 01-NOV-1996 (TREMBREL_01, Last sequence update)
DT 01-MAR-2001 (TREMBREL_16, Last annotation update)

DE INAD PROTEIN.
GN INAD.
OS Calliphora vicina (Blue blowfly) (Calliphora erythrocephala).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Oestroidea; Calliporidae; Calliphora.
OX NCBI_taxid=7373;
[1]
SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=CHALKY; TISSUE=RETINA;
RX MEDLINE=96216082; PubMed=8662634;
RA Huder A., Sander P., Paulsen R.;
RT "Phosphorylation of the Inad gene product, a photoreceptor membrane
protein required for recovery of visual excitation.";
RJ J. Biol. Chem. 271:11710-11717(1996).
CC -FUNCTION MAY BE INVOLVED IN CONTROL OF THE LIGHT RESPONSE. A RISE
IN INTRACELLULAR CALCIUM LEVELS UPON VISUAL EXCITATION MAY
INITIATE PHOSPHORYLATION OF THE INAD PROTEIN BY EYE-PC.

CC PHOSPHORYLATED INAD MAY IN TURN ACT ON ANOTHER PROTEIN SUCH AS TRP
 CC OR NORIA WHICH BOTH CO-PRECIPIRATE WITH INAD.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN RHADOMERAL PHOTORECEPTORS.
 CC -1- PFM: PHOSPHORYLATED BY EYE-PKC (PROBABLE).
 CC -1- SIMILARITY: TO DROSOPHILA INAD.
 DR EMBL: 269889; CAA93758.1; .
 DR HSSP: P31016; 1BFE.
 DR InterPro: IPR001478; .
 DR Pfam: PF00595; PDZ; 5.
 DR SMART; SM00228; PDZ; 1.
 DR Repeat: Phosphorylation; Glycoprotein.
 FT DOMAIN 41 420 2 X APPROXIMATE REPEATS.
 FT REPEAT 41 82 1.
 FT CARBOHYD 381 420 2.
 FT CARBOHYD 5 5 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT MOD_RES 15 15 BY PKC (POTENTIAL).
 FT MOD_RES 168 168 BY PKC (POTENTIAL).
 FT MOD_RES 185 185 BY PKC (POTENTIAL).
 FT MOD_RES 320 320 BY PKC (POTENTIAL).
 FT MOD_RES 321 321 BY PKC (POTENTIAL).
 FT MOD_RES 437 437 BY PKC (POTENTIAL).
 FT MOD_RES 595 595 BY PKC (POTENTIAL).
 FT MOD_RES 658 658 BY PKC (POTENTIAL).
 SQ SEQUENCE 665 AA; 73349 MW; C453192E18202BF7 CRC64;

Query Match 8.1%; Score 564.5; DB 5; Length 665;
 Best Local Similarity 26.8%; Pred. No. 1.7e-25;
 Matches 190; Conservative 109; Mismatches 229; Indels 181; Gaps 27;

OY 521 OPRVRLMREPSKSLGSIYVGRGMSRLSNGEVRGIFIKHVLHEDSPAGKNGTLKPRGD 580
 DB 10 QVQSVTLDRKTKKSFGLSIVRGEA-----RDGSKGIFIKGIYVPSPHGLGKIVGDR 64
 OY 581 IYBAPS---QSESEPE----- 593
 DB 65 LTLNKGVDVADTEPEVINLIKQAGSKIDLETQYGEQSNKMGNEIKENGESQNNMEN 124
 OY 594 ----KAPLCVSP-----PPPSAFAEMGSDHTOSSAS----- 621
 DB 125 QDSIKOPTKOPAIKPOOSKQAPGRPPVNMALNKSNTTSSKSDONLDDDEPTDRM 184
 OY 622 -----KISQDYDK--EDEFYSMKNIIRERYGLTGLHMIET 656
 DB 185 TGRIRTAAGEIDRASAGNCKLNKMEKDKETDEDEFGYTKAKINKRYTMNR-DLKKLEI 243
 OY 657 EK-GHSGLGLSLAGNKRBSMSVFIYIDPNC-AAGKDRGLQIADLELNGOILYGRSH 714
 DB 244 VRTNTALALAGHSDRCKMCFVAGVNTSGPLASVD--IKSGDEILEVNGVTLKNRCH 301
 OY 715 QNASSTIKCAPSKVILIFIRNKDAVNQMAVCPGNAVEPLPSNSENLOKNETEPTVTSDA 774
 DB 302 LNASVIFKINDE-KLVLTISRKKPND-----EGMSVKPI-----KKFPEI--DDT 345
 OY 775 AVDLSSFKNVHLELPKDDGGIAGI--SEEDTLGVIKSLTERGVAATDRGLKVGQ 831
 DB 346 KLFEGEYAKARASVQ--KKGEGFLGIMVYKGHVEVNGIGIFISDLRESNMMLAG-LAVGM 403
 OY 832 ILAVDEIYVYGIETKIFISLTKAKMTVKLTITHAENPDQOAVPSAAGAAGEKNSQSL 891
 DB 404 LLAVNNDVCSNYSYDAVALMLKRAEIVNLVLTLEETDI-----KKEEEE 452
 OY 892 WYPOGSP--EPESIRMTSSSTPAIFASDPATPPIPGCEETIEISKGRGLSLIVGGS 950
 DB 453 KKKEAARPEPER-----PVDSILEVYVGRKIYEVATDKKPLGIVYVGGK 499
 OY 951 DTLL--GATIHVEYEGAGCKDRLMAGDQILEVNGIDLKRAETHDAINVL----- 1001

DB 500 NNVTGCTIIFH-IYEGVIAEDKRLKIPDHIIQVNG-----KEVCEANTTLKHQLFYT 554
 OY 1002 QTPQRRRLLYDEAPYKKEEVCDDTLITLQKPKGGLSLYGRNDRGVSDIVKGG 1061
 DB 555 LYKATITQVYRADP-----EV-ETFKVEFARKAGDGLSLA--PNEGCTISSETTSAG 607
 OY 1062 IADPDRLIQDQIILLVNGEDVRNMSQEAVALKSLGTVLEVGRIK 1110
 DB 608 YADINKLQGRDITTFKNGDSLEGLTFEYCALFKGATGKISLEITRPK 656

RESULT 14
 O9NBV3 PRELIMINARY; PRT; 674 AA.
 AC O9NBV3.
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 GN INAD.
 OS INAD OR CG3504.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CANTON-S;
 RA Ashari M., Shieh B.;
 RT "The Inad locus in Drosophila."
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF245280; AAF81203.1; .
 DR Flybase; FBgn001263; Inad.
 DR InterPro; IPR00582; .
 DR InterPro; IPR001478; .
 DR Pfam; PF00595; PDZ; 5.
 DR Prodom; PD002965; .; 1.
 DR SMART; SM00228; PDZ; 1.
 SQ SEQUENCE 674 AA; 74305 MW; 55B1C57B7A100FID CRC64;

Query Match 8.0%; Score 556; DB 5; Length 674;
 Best Local Similarity 26.5%; Pred. No. 5.7e-25;
 Matches 202; Conservative 134; Mismatches 292; Indels 134; Gaps 31;

OY 368 VMLQNSKESFERTINIAKNSLSGTFVSANKDGLMIVRSIIGHGAIISRDRIAGDCI 427
 DB 18 VTLDRKTKKSF--GLCIVRGE-----VKDSPNTKTTGIFIKGIYVPSPHGLGKIVGDR 71
 OY 428 LSLNESSTISVYNAQRAAMLRLRHSILGPDIKITTYVAHELEFKISLSGOOSGRVMAIDF 487
 DB 72 LSLNGK---DVNRSTEQAVI--DLIKE-----ADFKIEL-----ELQTFD-- 106
 OY 488 SSYTGDIPELPEREGEGESEELQNTAYSNNOPRVELMREPSKSLGSIYVGRGMS 547
 DB 107 -----KDEQAKSDPRSNGYMAKKNFNEQOTNNNAS--GGGGMQ 147
 OY 548 RLSNGEVRMGIFIKHVLHEDSPAGKNGTLKPGDRIYBAPQSESEPEKAPLCVPPPPSA 607
 DB 148 GOGGOGGMAGM-----NRQOSMOKRNTTFTASMKQKSHNADEDEDEPTDRMTG-----RI 197
 OY 608 PLEMGSDHROSSA-----SKISQDYDK--EDEFYSMKNIIRERYGLTGLHMIETLEKH 660
 DB 198 RTBAGTIDIRASAGNCKLNKOEKDRKEDDEDEFGYTKAKINKRYTMNR-DLRIRIEYQRA 256
 OY 661 S-GIGLSLGNKDRBSMSVFIYIDPNGAAGKDRGLQIADLELNGOILYGRSHONASS 719
 DB 257 SKPLGLALAGHDKRQMACFVAGVDPNGALGSVD--IKPDDELVEVNGVTLKNRCHLMSA 315
 OY 720 IIKCAPSKVILIFIRNKDAVNQMAVCPGNAVEPLPSNSENLOKNETEPTVTSDAVIDLS 779
 DB 316 VRSYVDG--KLWLTISRKKPNDGMC-----VRFI-----KKFPAASDETKFI--FD 359

QY 780 SFKNVHLEPRDOGLGIAI---SEEDTSGVIKSLTEHGVAATDGRKVGDOILAVD 836
 DB 360 QEPKARIVQVRK-EGGILIMYIVYKHAIEVSGIFISDLREGSNMELAG-VKVGMDLAVN 417
 QY 837 DEIVYGVYIEKFIISLTKAKTKYK---LTHAENPDSQAVSAGAGAGEKNSQSILMV 893
 DB 418 QDVLESYWDATGTLKKAEEGVYMIITLTKSE---EAIKAEKAEKKEEAKKEEBEK 473
 QY 894 POSGPEPESIRNTSRSTPAIFASDPATCPILPCETITEISKRTGLSIVGSD-- 951
 DB 474 PQ-----EPATAEIKPKKLLILELKYKRMGYIVCGCKNNH 510
 QY 952 TLLGAFITHEVYEGSACKDGLNAGDOILEVNG--IDLKRAFDEALNVRQPQR-VR 1008
 DB 511 VTTGCVITH-VYPEGQVADRKLFKFDHICDINGTPHVGSMTELLKVLQPLFTTYEKAVT 569
 QY 1009 LTVIRDEAPYKEEBECDLTLELOKKPGKGLSIVGKRNTOGVPSIVGSIADPGR 1068
 DB 570 LTVFADPPELEK-----FVNDLKKKAKKEGLSL--SPNEIGCTIADLQIQVPEIDSK 622
 QY 1069 LIQGDQILLVNGEDVNRASQBAVALLKCSLGTVLEVGRIK 1110
 DB 623 LQRGIIIRKFNQDLEGLPFOVCYALFKGANGKSMETTRPK 664

RESULT 15
 Q24008 PRELIMINARY; PRT; 674 AA.
 AC Q24008
 DT 01-NOV-1996 (TREMblrel. 01. Created)
 DT 01-NOV-1998 (TREMblrel. 08. Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16. Last annotation update)
 DE INAD OR CG3504.
 GN EMBL OR CG3504.
 OS Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta:
 CC Pterygota: Neoptera: Endopterygota: Diptera: Brachycera: Muscomorpha;
 CC Ephydroidea: Drosophilidae: Drosophila.
 CC NCBI TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY:
 RA MEDLINE-20196006; PubMed-10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.C., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Bokora D., Botchan M.R., Bouck P., Brokstein P., Brotler P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Paulos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson D.R., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Jostali D., Houston K.A., Howland T.J., Wei M.-H., Ideyewam C.,
 RA Jostali M., Kaulush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclik J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheffler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.W., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster".
 RL Science 287:2185-2195(2000).
 RN [12]
 RP SEQUENCE FROM N.A.
 RC TISSUE-COMPOUND EYE.
 RA Lindsley D.L., Zimm G.G.,
 RT "The genome of Drosophila melanogaster".
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 RN [13]
 RP SEQUENCE FROM N.A.
 RC TISSUE-COMPOUND EYE.
 RA MEDLINE-95127229; PubMed-7826638;
 RA Shieh B.H., Niemeyer B.,
 RT "A novel protein encoded by the Inad gene regulates recovery of visual
 RL transduction in Drosophila".
 RL Neuron 14:201-210(1995).
 RN [14]
 RP SEQUENCE FROM N.A.
 RC TISSUE-COMPOUND EYE.
 RA Shieh B.,
 RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
 RN [15]
 RP SEQUENCE FROM N.A.
 RC TISSUE-COMPOUND EYE.
 RA Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003458; AAF6915.1;
 DR EMBL; U15803; AAC36490.1;
 DR HSSP; P31016; 1BFE.
 DR FlyBase; FBgn0001263; Inad.
 DR InterPro; IPR001478;
 DR Pfam; PF00595; PDZ; 5.
 DR SMART; SM00228; PDZ; 1.
 SQ SEQUENCE 674 AA; 74332 MW; DAC24091D9EA7E3 CRC64;

Query Match 7.98; Score 554; DB 5; Length 674;
 Best Local Similarity 26.58; Pred. No. 7.3e-25;
 Matches 202; Conservative 134; Mismatches 292; Indels 134; Gaps 31;

QY 368 VMLQNSKESERINTINAKSSLSGMTVSANKDGLGMYRSHIHGALSRGRIADGCI 427
 DB 18 VTLKTKGKSF--GICIVRGE-----VKDSPYTKTGTIFIKGIVDPSPAHLCGRKVGDR 71
 QY 428 LSINESTISVTNQAARMLRRHSLIGPDIKITYVPAHELEFKISLQOQGRVMAIDIF 487
 DB 72 LSLNGK---DVNRNTEQAVI---DLIKE-----ADRIEL-----ELQTFD-- 106
 QY 488 SSYGRDIPELPEREEGEGESELQNTAVSNMOPRRVYELMRPESKLSIGISVGRMG 547
 DB 107 -----KSDQQAQSPDRSNGYQAQKKNFQEOQTNNAS--GGQGMQ 147
 QY 548 RLSNGEVRGIFIKHVLDEDSFAGKNGTLKPRDRIYEAPOSSEPERAPLCSYVPPPPSA 607
 DB 148 GQGGGQAGAG---NRQOSQKRNRTFTTASMRKSHSYADEDEDTRDMWG-----RI 197
 QY 608 FAEGSGDHDTOSA-----SKTSDPVR--EDEFYSKKNIRERYGTLGLHMLELEK 660
 DB 198 RTEAGYELDRASAGCNCLNKOEDKRDKEDEFEFTYAKIKRNNMK--DLARLEVG 256
 QY 661 S-GIGLSLAGKDRSRMSVFIVGIDPNGAAGKDRLOJADELLEINQILYGRSHQNAS 719
 DB 257 SKPLGLALAGHKDRQAKAFVAGVDPNGALGSVD--IKRDEIYVANGNVILKNRCHLNASA 315
 QY 720 IIRKAPSKVITIFIRKNDVAVQMAVCGNAVEPLPSNSENLQNKETEPYVTSAAVDLS 779
 DB 316 VFKNVGD-KLVMTTSRRKPNDEGMC---VKPI-----KKPPTASDETKEFI--FD 359


```

OY 780 SFKNVQHLELPKDOGGLGIAI---SEEDTLSGVIIKSLTEHGVAAATDGRKVGDOILAVD 836
Db 360 QPFRKFTVQVRK-EGFLGIMVIYGHAEVSGSIFISDLREGSNAELAG-VKVGDMLLAVN 417
OY 837 DEIVGVPIEKFTISLTKTAKMTVK---LTIHAENPDQAVPSAGAASGKKNSQSILMV 893
Db 418 QDVTLIESNVDQATGLKRAEGVVTMLTLTKE---EAIKAKAAEKKKEAKKEEK 473
OY 894 PQSGSPEPESIRNTSRSSPALFASDPATCPIIPGCEFTTIEISKGRGTGLSLIVGSD-- 951
Db 474 PQ-----EPATAETKPNKILIELKVEKKPMGVIVCGKNNH 510
OY 952 TLLGAFIIEVYEBGACDGRMLNAGDOILEVNG--IDLKATHDEAINVLRQTPQR-VR 1008
Db 511 VTTGCVITH-VYPEGOVADKRLKIFDHICDINGTPIHVGSMTTLKVHQLFHTYEKAVT 569
OY 1009 LTLVHDEAPYKEEVEYCDTLTIELQKPGKGLSIVGKRNDTGVFVSDIVKGIADPDGR 1068
Db 570 LTVFRADPELEK-----FNVDLMKKAKGELGLSL--SPNEIGCTIADLIQGYPEIDSK 622
OY 1069 LIQSDQILLVNGEDVYRNASQEAVALLKCSLGTVTEVGRK 1110
Db 623 LQRCDIITKFNGBALEGLPFQVCVALFKGANGKVSMEVTRPK 664

```

Search completed: July 12, 2001, 14:43:01
Job time: 214 sec

•
•
•
•

W

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2001, 14:37:37 ; Search time 23.36 seconds

(without alignments)
1184.015 Million cell updates/sec

Title: US-09-502-698-1
Perfect score: 6990
Sequence: 1 MCCRRTVPTTQSELDLD.....EAAVALIKRTKGYTLNVL 1373

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	509	7.3	2485	US-09-290-640-46	Sequence 46, Appl
2	500.5	7.2	2466	US-09-080-855-12	Sequence 12, Appl
3	500.5	7.2	2466	PCT-US94-09943-2	Sequence 2, Appl
4	498	7.1	2465	US-08-596-291-3	Sequence 3, Appl
5	498	7.1	2465	US-09-100-804-3	Sequence 3, Appl
6	340	4.9	1112	US-09-045-632-2	Sequence 2, Appl
7	340	4.9	1112	US-09-045-632-3	Sequence 3, Appl
8	335	4.8	1018	US-09-045-632-15	Sequence 15, Appl
9	335	4.8	1018	US-09-045-632-32	Sequence 16, Appl
10	335	4.8	1061	US-08-410-804-1	Sequence 32, Appl
11	334.5	4.8	610	US-08-259-514-1	Sequence 1, Appl
12	334.5	4.8	610	US-08-858-311-1	Sequence 1, Appl
13	334.5	4.8	610	US-08-858-311-1	Sequence 1, Appl
14	334.5	4.8	1050	US-09-045-632-50	Sequence 50, Appl
15	331.5	4.7	1050	US-09-045-632-49	Sequence 49, Appl
16	329.5	4.7	604	US-09-045-632-14	Sequence 14, Appl
17	327	4.7	918	US-09-045-632-21	Sequence 21, Appl
18	327	4.7	961	US-09-045-632-33	Sequence 33, Appl
19	321	4.6	602	US-09-045-632-20	Sequence 20, Appl
20	288.5	4.1	861	US-09-045-632-25	Sequence 25, Appl
21	287	4.1	818	US-09-045-632-13	Sequence 13, Appl
22	273	3.9	504	US-09-045-632-19	Sequence 19, Appl
23	271	3.9	504	US-09-045-632-35	Sequence 35, Appl
24	266.5	3.8	642	US-09-045-632-28	Sequence 28, Appl
25	265	3.8	599	US-09-045-632-24	Sequence 24, Appl
26	258.5	3.7	502	US-09-045-632-24	Sequence 24, Appl
27	252	3.6	374	US-09-091-405-2	Sequence 2, Appl

28	227	3.2	519	3	US-08-997-445D-2	Sequence 2, Appl
29	225.5	3.2	488	3	US-09-045-632-27	Sequence 27, Appl
30	219	3.1	293	3	US-09-045-632-30	Sequence 30, Appl
31	219	3.1	541	3	US-09-045-632-36	Sequence 36, Appl
32	217	3.1	284	3	US-09-045-632-12	Sequence 12, Appl
33	214.5	3.1	407	3	US-09-045-632-18	Sequence 18, Appl
34	209.5	3.0	1612	3	US-08-545-860D-48	Sequence 48, Appl
35	209.5	3.0	1612	5	PCT-US94-04496-48	Sequence 48, Appl
36	203	2.9	404	3	US-09-045-632-23	Sequence 23, Appl
37	200	2.9	79	3	US-09-100-804-27	Sequence 27, Appl
38	200	2.9	86	3	US-08-545-860D-53	Sequence 53, Appl
39	200	2.9	86	5	PCT-US94-04496-53	Sequence 53, Appl
40	196.5	2.8	198	3	US-09-045-632-11	Sequence 11, Appl
41	190.5	2.7	1829	4	US-09-157-420-11	Sequence 1, Appl
42	178.5	2.6	182	3	US-09-045-632-29	Sequence 29, Appl
43	176.5	2.5	233	2	US-09-151-611-1	Sequence 1, Appl
44	173	2.5	505	1	US-08-123-161A-14	Sequence 14, Appl
45	173	2.5	505	1	US-08-483-278-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-290-640-46
Sequence 46, Application US/09290640
Patent No. 6204055
GENERAL INFORMATION:
APPLICANT: Dean, Nicholas M.
TITLE OF INVENTION: Anti-sense Compound Modulation of Fas Mediated Signaling
FILE REFERENCE: ISPH-0351
CURRENT APPLICATION NUMBER: US/09/290,640
CURRENT FILING DATE: 1999-04-12
NUMBER OF SEQ ID NOS: 85
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 46
LENGTH: 2485
TYPE: PRT
ORGANISM: Homo sapiens
US-09-290-640-46

Query Match 7.3%; Score 509; DB 4; Length 2485;
Best Local Similarity 22.8%; Pred No. 36-37;
Matches 243; Conservative 140; Mismatches 371; Indels 314; Gaps 37;

QY 530 EPEKSLGISVGGGSGRLSNGEVMKGIPIKHLDESPAGKNGTLKPGDRIVEAPSOSE 589
DB 990 EPPPTVAELVGRKPSHMSRSDAELAGV---TKLNKSVASLNRSPERRKHSSSSSI 1046
QY 590 SEPEKAPLCSVPPPPSAFEMGSDHTQSSAKISQDVDEDFGYSWKNIRRYGTLLG 649
DB 1047 EDEGAYVVGMT-----MHSSGNSSQVPLKENDV-----LHRMSIVSS 1086
QY 650 ---ELMIELEK-GHSGIGLSLGNKDRSR--MSVFTVIGIDPGANGKGRQIADELE 703
DB 1087 PEREITLVNKKAKYGLGQITIGGEMGRLDGITISSVAPGPADLDGCLPGRLIS 1146
QY 704 INQILYGRSHQVASSIIICAPSKVRIIFRNKDNOMAVCGNAVEPLPSNENLQNK 763
DB 1147 VNSVSLGVSHNAIETLQNAPEDEVTVLSQPEKISKVSTVH---LTNEMKVMYMK 1202
QY 764 ETEPTVTSDAVDSLFKN-----VOHLELPK--DQGL--GIAISEEDTISGVIYS 813
DB 1203 SS---YMODSAIDSSSKDHMSRGLRHTISENSFPGSLRGSSSDPSRESASLSQ 1258
QY 814 LTFHGVAATGRKLVQDQILAVDDEVYGVPIKFTSLKTAKMTYKTLTHAENPSONAV 873
DB 1259 SQVNGFPAS---HLDDQTV---QESOHGSPSPSVIS---KATEKETDSNQSKTK 1306
QY 874 PSAAGASGEKK-----NSSOSLWPGSGSPESPEIRTSRSRSPAIRPASPPATC 923

Db 1845 ---VIGRVELPRIML-----PHLLPITLTCKNEELGFSICGHDLSL 1885
Qy 1312 HDDLPIYKTVAKGASDGLKRGDOLIANVNGSLBEVTHEA-----VAILKR 1362
Db 1886 Y--QVYISDIPRSVAALIEGMLQDLIDVHYNGVSTOQMTLEEVNRLALDMSLPSLVLKA 1943
Qy 1363 TKGVTLM 1370
Db 1944 TRNDLPV 1951

RESULT 3
PCT-US94-09943-2

Sequence 2, Application PC/TUS9409943
GENERAL INFORMATION:
APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL
TITLE OF INVENTION: EXPRESSION OF NUCLEOTIDE SEQUENCES FOR NOVEL PROTEIN
TITLE OF INVENTION: TYROSINE PHOSPHATASES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
STREET: 600 ATLANTIC AVENUE
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09943
FILING DATE: 01-SEP-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/115,573
FILING DATE: 01-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: TWOMEY, MICHAEL J.
REGISTRATION NUMBER: P-38,349
REFERENCE/DOCKET NUMBER: L0461/7000WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/720-3500
TELEFAX: 617/720-2441
TELEX: 92-1742 EZEKIEL
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2466 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-09943-2

Query Match 7.2%; Score 500.5; DB 5; Length 2466;
Best Local Similarity 22.2%; Pred. No. 1.5e-31;
Matches 255; Conservative 146; Mismatches 396; Indels 351; Gaps 41;

Qy 510 ELQNTAISMNMPRVELMREPSLSIGISIVGGRMGSKLSNG----- 552
Db 868 QLOMRAROSNODADODIERASFRSLNLQASVIRGFNNGRAISTGSLASSTLKLAVPLISV 927
Qy 553 --EVARGIFIKHV-----LEDSPAGKNGTL--RPGD--RIVEAPSOSESEPEKAPLCS 599
Db 928 QAEILKRLSCSLSLIXPLQNSSKEKNKASWEKPREMSKSYHLSQASLPHKKNVIV 987
Qy 600 VPPPPSAFAEN-GSDHTQSSAS-----KISQDVKED 631

Db 988 NMEPPQVAVELVGRKPSHQMSRSDAELAGVTYTKLNNSKSVASLNSPERKHHESDSSIE 1047
Qy 632 ERG--YSKNNIERKGTLTG---ELHMLEK-GHSGGLSLAGNKDRSR--MSFIVCI 683
Db 1048 DPGQAVLDLKHRSYSSPEREITLVNLKDKARGLAFOIIGCKMRDLGIFISSV 1107
Qy 684 DNGAAGKDRGROIADDELEINGOILYGRSHONASSIICAPSKYIIFIRKNDANOMA 743
Db 1108 AFGGRADFHGCLKPGDRLISVNSVSLGVSHHALETIONAEDDTLVISQREKISKVP 1167
Qy 744 VCPGNAVEPLPSNSENLOKETEPTVTSDAAVDLSSFFKN-----VOHLEPK--DOGG 795
Db 1168 STPVH---LTNEMKNYMKSS---YMODSAIDSSKDHMSRGLRHISENSFCPSGG 1219
Qy 796 L--GLAISEEDLSGVIITSLTEHGVAAIDGRLKRGDOLAVDELIVGPIEKFISLKL 853
Db 1220 LREGSLSDSRTESASLSISOVNGFFAS---HLGDQW---QESQHSPPSPVYS-- 1269
Qy 854 TAKMTVKLTIAENPDSQAVPSAAGAAGEKK-----NSSQSLAVPQSGSPSPS 903
Db 1270 --KATEKETFTDSNOSKTKKPGISDVTYSDRGDSMDAATYSSSDHOTPKO---ESS 1324
Qy 904 INTRKSTPAIFASDPATCPIPGCETIIEISKRTGIGLSIVGSDTL--GAFIIE 961
Db 1325 SVNTSKMKNKTFESSPPK---PGDIFEVELAKNDNSIGIVTGVTNVRHGGIYVKA 1380
Qy 962 YVEGACAKDGLMAGDOLLEVNGIDLKRAIHDEALINVRQTPORVRLTYDEAPYKE 1021
Db 1381 VIPOGAAESDGRIRHKGDRVLAVNGVSLGATHKQAVETIRNGOVVHLLLEGQSPSTKE 1440
Qy 1022 EV---C-----DTLIELOKKRGKIGLS-- 1042
Db 1441 HVPVTPQCLTSQONAGCPPEVKTKTQVKDYSPYTEETFEKFLKN--SSGLGFSFSRE 1499
Qy 1043 ---IVGKRNDSVFPVSDIYKGIADPDRLLIOGDOILLVNGEDVNRASQEAVALKCSL 1099
Db 1500 DNLIEQINASIVRVKIKLFAQGPAAESGKIDVDYILKYNKASLNGISQGEVYSALRGTA 1559
Qy 1100 GTVTELEVGRIRKG-----PHESE- 1117
Db 1560 PEVFLILCRPPGVLPLEIDTALLPLQSPAQVLPNSSKDSQSPCEVOSTSDENEMSDK 1619
Qy 1118 ---RRPQTSQVSEGS-----LSPTF- 1138
Db 1620 SKKQCKSPSRSDSDSSGSDDLVTAPANISNTWSALSALHQTLSNMYSQAOSSHHAEPK 1679
Qy 1139 SGGST-----SSLESSK- 1152
Db 1680 SOEDTICTMFWYPRKIPNKPFEEDSNPSPLPPDMAFGQSYOPQSEASSSMDKYHIHI 1739
Qy 1153 ---KNALASITQGLR---TYEMKKGPTDSLSTIAG--VGPLGDPV 1192
Db 1740 SEPTROENMTPKNDLENLEDEFELEVELLITLISEKASLSLFTYKGNORIGCVYHVD- 1798
Qy 1193 IFIAMHPTGVAOQOKLEVRGRIYICTSTEGMHTQVDAVNLKNASISIEQVAVAGD 1252
Db 1799 ---IOP---AKSDGRKLPGRDLIKVNDTDTNMHTDAVNLRAASTVNL----- 1844
Qy 1253 VSVVTGHHQE-PASSLSLFTGLTSTISIFODDLGPPCKSITLERGPDGIFGIVGCGSP 1311
Db 1845 ---VIGRVELPRIML-----PHLLPITLTCKNEELGFSICGHDLSL 1885
Qy 1312 HDDLPIYKTVAKGASDGLKRGDOLIANVNGSLBEVTHEA-----VAILKR 1362
Db 1886 Y--QVYISDIPRSVAALIEGMLQDLIDVHYNGVSTOQMTLEEVNRLALDMSLPSLVLKA 1943
Qy 1363 TKGVTLM 1370
Db 1944 TRNDLPV 1951

RESULT 4
US-08-596-291-3

Db 1168 TPVH---LTNEMKNVKKSS---YMDASLIDSSSKDHNMWNGTLRIHSENFGPSGGI 1219
 QY 797 ---GLAISEDILSGYIISKLTETHGVAATDGRKLVGDQILAANDELYVGPRIEKFISLKT 854
 Db 1220 REGSLSSODSRSTESASLSOSQVNGFFAS---HLGQTV---QESOHGSPSPSVIS--- 1268
 QY 855 AKMYTKLTHAENPDSQAVPSAAGAASGEKK-----NSSQGLAMPQSSPEPEST 904
 Db 1269 -KATKEKEFTDSNOSKTKRKPGISDVTDYDRDSDMDDEATYSSSDQHGPVK---ESSS 1324
 QY 905 RNTSSSTPAIFASDPATCPPIIPGCEITIEISKRTGLSLIVGSGSDTL--GAFIHEX 962
 Db 1325 VNTSKMNFKEFFSSPPK---PGDIFEWELKKNDSLISVYTGAVNTSVRHGIVYKGV 1380
 QY 963 YEEGASCKDGRFAMADQILEVNGIDLRKHTHDEAIVLWTFOPORVLTLYRDAFAPKEE 1022
 Db 1381 IPQGAESDGRIRHKDRLVLANVGSLEGAITHQAVETLANTQGVHLLLEKQSPSPSKEH 1440
 QY 1023 V-----C-----DTLTLELQKPKGIGLS----- 1042
 Db 1441 VPVTPQCTLSDNMAOGGGEKVKTTQVKDYSFVTEENFPEVLEKFN--SSGLFFSFRSD 1499
 QY 1043 --IYVKRNDTGVEVDIYKGIADBDGRILQDQILLVNGEDVRNKSQEAVALLECSLG 1100
 Db 1500 NLIPQIINASTYRVKRLPAGQPAAESGKIDVGDYILKVNAGSLKGLSQDEYISALNGTAP 1559
 QY 1101 TVTLEVGRIKAG-----PFHS----- 1117
 Db 1560 EYFLLCLRPPLVELEIDTALLTPLOSFAQVLPNSSKDSQSPCEQSTSDENEMSDKS 1619
 QY 1118 ---RRPSQTSQVSPGS-----LSSFP-----PKS 1139
 Db 1620 KKQCKSPSKRSDYSOSSGSGEDDLVTAPANTISNTWSSALHOTLSMVGQASHHEAPKS 1679
 QY 1140 GSST-----SESLSSSK----- 1152
 Db 1680 QEDTICTMFEYPOKIRPKREFEDSNPSRLPRDPAQSQYQOPSESASSSMOKYTHIHIS 1739
 QY 1153 -----KNALASELOGLR-----TVEMKKGFTDSLGISIAG---VSPLODVP 1193
 Db 1740 EPTROENMTPRLNDENLHDELEVEELLITTIKSEKALGFTVTGNORIGCYADV- 1797
 QY 1194 FTAAMHPGVAQOTOKLRVGDRIYVTCGSTGEMTHTOAVNLKNAAGSIEM-----OV 1247
 Db 1798 ---IODP---AKSDGRKAKGRDLIKVNDVYNMHTDAVNLRAASKYRVLVIGVPR 1851
 QY 1248 VAGDV-SVYTGHHOEPASSLSLFTGLNSTSLFODDLGPQCKSTILERGPDGLGFSIVG 1306
 Db 1852 TQNTNVASFPATG-----KLTGNKEBELGSLG 1879
 QY 1307 GYGSPHGDLPIYKTVFKAKGASEDGRKRGQIILVANGOSLEGVTHEEA-----V 1357
 Db 1880 GHDSLY--QVYISDINPRSAVAIEGNLOLDLIVYNGVSTQGMLEEVNRALDMSLPS 1937
 QY 1358 AILKRTKGTITLM 1370
 Db 1938 LVLKATRNDLPVY 1950
 RESULT 5
 US-09-100-804-3
 : Sequence 3, Application US/09100804
 : Patent No. 6066472
 : GENERAL INFORMATION:
 : APPLICANT: GONEZ, LEONEL JORGE
 : APPLICANT: SARAS, JAN
 : APPLICANT: CLAESSON-WEIS, LENA
 : APPLICANT: HELDIN, CARL-HENRIK
 : TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL
 : TITLE OF INVENTION: EXPRESSION OF NUCLEOTIDE SEQUENCES FOR NOVEL PROTEIN
 : TITLE OF INVENTION: TYROSINE PHOSPHATASES
 : NUMBER OF SEQUENCES: 34
 : CORRESPONDENCE ADDRESSES:
 :

ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
 STREET: 600 ATLANTIC AVENUE
 CITY: BOSTON
 STATE: MASSACHUSETTS
 COUNTRY: USA
 ZIP: 02210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/100,804
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/596,291
 FILING DATE: 09-AUG-1996
 APPLICATION NUMBER: US 08/115,573
 FILING DATE: 01-SEP-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/09943
 FILING DATE: 01-SEP-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: GATES, EDWARD R.
 REGISTRATION NUMBER: 31,616
 REFERENCE/DOCKET NUMBER: L0461/7003
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441
 TELEX:
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2465 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-100-804-3

Query Match 7.1%; Score 498; DB 3; Length 2465;
 Best Local Similarity 21.9%; Pred. No. 2,4e-31;
 Matches 252; Conservative 150; Mismatches 389; Indels 362; Gaps 41;

510 ELONATAYMNMNRRVELMREPSKSLGISTVGRMGSLNSG----- 552
 868 QLOMRAROSNODADIERASFRSLNLQAESVGFNMGRALISTGLASSTLNTKLAVPLISV 927
 553 --EVARGIFIKHV-----LEDSPAGKNGTL---KPGD--RIVEAPSOSESPEKAPLCS 599
 928 QAEILKRLSCSLSLYQPLQNSKSKKNDKASWEKPREMSKSYHDLQASLYPHKKNIV 987
 600 VPPPPSAFAEM-GSDHTOSSAS-----KTSQDVKED 631
 988 NMPEPPQVYAEIVGRKPSHQMSRSDAESLAGVTKLNNSKSVASLNSRPERKHESDSSSIE 1047
 632 EFG--YSKNKNIERYCITLG---ELHMIETELR-GHSGLGLS-LAGNKRDSRSVITVGD 684
 1048 DPGQAYVLVDLHKRWISYSPERETTLVNLKDKARYGLGFOIIGGEKMETDGIETISSVA 1107
 685 PNGAGKGDRLQIADLELLEINGOILYGRSHONASSIILKAPSKVILFIIRNDVANOVA 744
 1108 PGGADPFHGCCLKPGRDLISVNSVLEGVSHAAIEILQNAFPDVLIVLSQPEKISKVPS 1167
 745 CGNAVEPLPSNENLONKETEPTVTTSDAVALDSSFKN-----VQHLLEPK--DOGGL 796
 1168 TPVH-----LTNEMKRYMKKS---YMODSAIDSSSKDHMGRTLRHSSENSFGSGGL 1219
 797 --GIAISEEDTLSTVILSLTEHGVAAITDGRKLKVDQILAVNDELIVGPIPIKFIISLKT 854
 1220 REGSISSDQSRSTESASLSOSQVNGFFAS---HLGDQTV---QESOHGSPSPSVIS--- 1268
 855 AKMTVKLTTHAENPDQAVPSPAAGAAGEKK-----NSSOSLWVPGSGSPEPESTI 904

1269 --KATEKETFTDSNOSKTKKPGISDVTDYSDRGDSMDERTYSSSDPHOTPKO---ESSSS 1324
 905 RNTSRSSTPALPASPDPATCPPIIPGCEETIEISKRGTLGSLIVGSDTL--GAFIIEHV 962
 1335 VNTSNKMNKFTSSSPK---PGDIFEVELAKNONSIGISTGVGVNTSVRRGCIYVADY 1380
 963 YEEGAACKDGRILMAGDQILEVNGIDLRKATHDEAINVLKQTPQRYRLTYRDEAPYKEE 1022
 1381 IFOGAESDGRITHKGDVILAVNGVSLGATHKQAVETLRNTGOVYHLLLEKQSPSTKEH 1440
 1023 V-----C-----DTLELQKKRGKGLS----- 1042
 1441 VFTPTQCTISDNOAGOGPEKKYKTTQYKDYSPVTEENTFEVKLRN--SSGLGFSFRED 1499
 1043 --TVGRNDTVGVFDIYKGIADPDGRLIQDQILLVNGEDVRNASQEAVALLKCSLG 1100
 1500 NLIPIQINASIVRYKKLFAGQPAESGKIDVDVLTKVNGASLKGSLSQOEVIASALRGAP 1559
 1101 TYTLEVGRIKAG-----PFHSF----- 1117
 1560 EVFLLICRPPGVLPLEIDTALITPLQSPAQVLPNSKSDSOPSCVEQSTSDENEMSDKS 1619
 1118 ---RRPSQTSQVSEGS-----LSSFTF-----PLS 1139
 1620 KQCKSPSRKDSYSDSSGSGEDDLVTAPANISNTWSSALHOTLSNMVSOAOSHHEAPKS 1679
 1140 GSST-----SELSSESSK----- 1152
 1680 QEDTICTMFTYPOKIPNKPEFEDSNPRLPRDMARGQSYQPQSEASSSMKYHIHHS 1739
 1153 ---KNALASEIQGLR---TVENKKRPTSLGISTINGG---VSPGLADPVI 1193
 1740 EPTROENMTPRLNDENLHDELEVELLITLIKESKASLGFTYVKGMRICCYHVDV-- 1797
 1194 FIAMHPGTVAAQOTOKLRYGDRIVTICSTEGMHTGOAVNLKNASGSIEM-----QV 1247
 1798 ---IDP---AKSDGRILKRGDLIKVNDIVNMHTDAVNLBRAAKTVRLVIGRVRI 1851
 1248 VAGDV-SVYTGSHQOEPASSSLFTGLSTSTIFODDLGPPCKSITLERPGDLGFSYVG 1306
 1852 TONTNVASFATCH-----KLNCNKEELGFSLCG 1879
 1307 GYGSPHGLPIYVKTVPKGAASEDEGRILKRGDQILAVNGOSLEGVTHEA-----V 1357
 1880 GHDSL--QVVTISDINPRSAIEGNLQLDLVITYVNGVSTQGMTLEVRNALDMSLPS 1937
 1358 AILKRTKGTITLM 1370
 1938 LVLKATRNLDLPVY 1950

RESULT 6
 US-09-045-632-2
 Sequence 2, Application US/09045632
 Patent No. 6001575
 GENERAL INFORMATION:
 APPLICANT: HUGANIT, Richard L.
 APPLICANT: DONG, Hualing
 TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
 TITLE OF INVENTION: GRIP-RELATED MOLECULES
 NUMBER OF SEQUENCES: 105
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
 STREET: 130 Water Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,632
FILING DATE: 19-MAR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,016
FILING DATE: 19-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Corleiss, Peter F.
REGISTRATION NUMBER: 33,860
REFERENCE/DOCKET NUMBER: 48147/1699-CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1112 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-045-632-2

```

```

Query Match      4.98; Score 340; DB 3; Length 1112;
Best Local Similarity 22.6%; Pred. No. 6,4e-19;
Matches 173; Conservative 90; Mismatches 265; Indels 238; Gaps 22;

```

```

OY 781 FKNVHLEPRDQ-GGIAISEDTLSGYI-----IKSLTEHGVAAATGRKAVGDOI 832
DB 48 FKGSTVELMKREKSTYLGK-----TVSGGIDKDKRPVSNLRGGIAASDOLDVGYI 101
OY 833 LAVDEIVGVPIEKFTSLTKAKTVKLTIAENPDQAVPSAAGAAGEKKNSSQSLM 892
DB 102 KAVGGINLAKFRHDEILSLKNGERVYLEVEYELP----- 137
OY 893 VPQSGSPESIRNTSSSTPAIFASDPATCPIIPGCTTIEISKRTG--LGLSIYGS 950
DB 138 -----PVSIQSS-----VMF-----RTVEVTLHKEGNTFGFVIRGCA 170
OY 951 D---TLGATFHEVYEGACAKDGRIMADQDILEVNGIDLRKATHEATINVLKOTQPV 1007
DB 171 HDRNRKSPVVTTCVRPGDPDRREGTTPKGRLLSVGIRLLGTHAEAMSLKQCGEA 230
OY 1008 RLTL-YRDEAPYKKEEVCDDTLFIELOKPKGKGLSIYGV--RNDTGVSVDIVKGIAD 1064
DB 231 TLLEIVDSANDSVATASGPLLEVAKTPGASLGVALTTSVCKKQYIVIDKISASIA 290
OY 1065 PDGRLLIGDQDILLVNGEDVRNASEAVAAALIKSLGIVTLEF-----GRK-AGPFH-- 1115
DB 291 RCGALHYGDHILSIDTSMETCYLAETQFLGNTTDQVKELIPHOIRLALKGPDHVKI 350
OY 1116 -----SERPSQ 1122
DB 351 QRSRQLEPMPWMASSQCSVHTNNHHNPHDHCVPALGFKALTPNPSPMWSSSSSTS 410
OY 1123 TSOVSEGLSSTFPPLSGSSSTS-----ESLSSSKNALASEIQTGRTV 1166
DB 411 MSAVSLSLNNGTLPRLSYTSPRGTMWRRLKKKDFKSSLSLSTAGLGGVHHETT 470
OY 1167 E--KKGPDSLGISINGV--GSPLDGVPFTIAMMPTGVAAGTOKLGVDRIVTIGTS 1223
DB 471 EYVLTADPVYGFGLQIGSVFATETLSSPLISYLEADSPERCGVLIQIGRVAAINGIP 530
OY 1224 TEGTHTQAVVLLKNASGS-----TEMVYAGGDSVYVGHQEPASSSLSFTGLSTS 1277
DB 531 TEDSTFEANOLDKRDSSITSKVTLIEFEDVAESVIPSSTGFHVAKLPKKHVEL-GITISS 589
OY 1278 IFDDDLGP-----QCK--- 1289
DB 590 PSSRKPDPLVTSIDIKKGSVAHRTGTLELGDKLAIINIRLSDCSMEDAVQIILQCCEDLV 649

```

```

OY 1290 -----SITLER--GPDGLGFSIVGYGSPHGLDPIYKTVFAK 1325
DB 650 KLIKRDENSDNEQESSGAIYVELKRYGGPLGITIS-----GTEPFDPIILISLTKG 704
OY 1326 GAASEDRKLRGDDIITAVNGSLGCVTHEEVALKRTKGTVMV 1371
DB 705 GLAERTGAIHIGDRILAIINSSSLKGPLSEDIHLLQMAGETVTLKI 750

```

RESULT 7

```

US-09-045-632-3
Sequence 3, Application US/09045632
Patent No. 6001575
GENERAL INFORMATION:
APPLICANT: Huganir, Richard L.
TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
TITLE OF INVENTION: GRIP-RELATED MOLECULES
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,632
FILING DATE: 19-MAR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,016
FILING DATE: 19-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Corleiss, Peter F.
REGISTRATION NUMBER: 33,860
REFERENCE/DOCKET NUMBER: 48147/1699-CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1112 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-045-632-3

```

```

Query Match      4.98; Score 340; DB 3; Length 1112;
Best Local Similarity 22.6%; Pred. No. 6,4e-19;
Matches 173; Conservative 90; Mismatches 265; Indels 238; Gaps 22;

```

```

OY 781 FKNVHLEPRDQ-GGIAISEDTLSGYI-----IKSLTEHGVAAATGRKAVGDOI 832
DB 48 FKGSTVELMKREKSTYLGK-----TVSGGIDKDKRPVSNLRGGIAASDOLDVGYI 101
OY 833 LAVDEIVGVPIEKFTSLTKAKTVKLTIAENPDQAVPSAAGAAGEKKNSSQSLM 892
DB 102 KAVGGINLAKFRHDEILSLKNGERVYLEVEYELP----- 137
OY 893 VPQSGSPESIRNTSSSTPAIFASDPATCPIIPGCTTIEISKRTG--LGLSIYGS 950
DB 138 -----PVSIQSS-----VMF-----RTVEVTLHKEGNTFGFVIRGCA 170
OY 951 D---TLGATFHEVYEGACAKDGRIMADQDILEVNGIDLRKATHEATINVLKOTQPV 1007

```



```

Db 171 HDDRNRKRVVITCVPRGGPDDREGTKRPGDRLLSVDIRLLGTTTHAEMSLKQCGQEA 230
Qy 1008 RLTL-YRDPAPKEEVCTLTIELOKKRGKLGISYVK--RNDTGVSIVKGIAD 1064
Db 231 TLLIEDVAMSVAATASCPRLVEVAKTPGASLGVALTTSVCCNQVYIDIKSASIA 290
Qy 1065 PGRLIOGQIILVNGEDVRNASEAVALLCSLGTVLEV-----GRIK-AGPFI--- 1115
Db 291 RCGALHVDHILSIDGTSMEYCTLAETQFLGNTDQVLELPHHQTLALKGPDPHVI 350
Qy 1116 -----SERRPSQ 1122
Db 351 QRSQDRLPMDPWASSQSVHTNHNHPHDPCHVPALGFPKALTPNSPRAWSSSSPTS 410
Qy 1123 TQOVSEGSISFTPLSGSSTS-----ESLESSKKNALASEIOGLRTV 1166
Db 411 MSAVSSLSINMGTLPRLSTSPRGTMRRRLKKDKFSSLSLASTVGLAQVHTET 470
Qy 1167 E--MKKGPDSIGISYAGV--GSPLDGVPIFTAMMPTGVAOTOKLRVGRIVTIGTS 1223
Db 471 EYVLRADPTGFGIOLOGSVFATELISPLISYIFADSPAERCGLVQIGDVMALNGIP 530
Qy 1224 TEGMTHQAVNLKNAAGS-----TEMQVAGDVSVYTGHHQBPASSLSFTGLTSTS 1277
Db 531 TEDSTFEENQOLLRDSITSKYTLIEFDVAVESVIPSSTFHVKLKPKHSVEL-GITISS 589
Qy 1278 IQDDLGP-----OCK--- 1289
Db 590 PSSRRPGDLVISDIKKSVAHRTGTLELDKLLAIDNIRLDSCEMADAVQILOQCEDEV 649
Qy 1290 -----SITLER--GPDGLGFSIVGSGPHGLPIYVTKVFAK 1325
Db 650 KIKIRKEDNSDEQESSGAIITYVELKRYGGLGITIS-----GTEPPDPPIISSLTNG 704
Qy 1326 GAASEDRKLRKGDQIIVANGOSLEGVTHEEAVAILKRTKGYTLMV 1371
Db 705 GLAERTGAIHGDRILAINSSSLKGRPLSEDIHLLQMAGETVTLKI 750

RESULT 8
US-09-045-632-15
; Sequence 15, Application US/09045632
; Patent No. 6001575
; GENERAL INFORMATION:
; APPLICANT: Huganir, Richard L.
; TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
; TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
; NUMBER OF SEQUENCES: 105
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/045,632
; FILING DATE: 19-MAR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/041,016
; FILING DATE: 19-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Cortless, Peter F.
; REGISTRATION NUMBER: 33,860
; REFERENCE/DOCKET NUMBER: 4817/1699-CIP
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 702 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-045-632-15

Query Match 4.8%; Score 335; DB 3; Length 702;
Best Local Similarity 22.5%; Pred No. 7.3e-19;
Matches 171; Conservative 90; Mismatches 261; Indels 238; Gaps 22;

Qy 787 LELPKDQG-GLGIAISEEDTLGVI-----IKSLTEHVATDRLKLVGQIILAVDE 838
Db 3 VELMKKEGTTLC-----TVSGGIDKDGKPRVSNLRGGIARSDQDLVDGYIKAVNGI 56
Qy 839 IYVGPTEIEFISLTKAKTVKLTTHAENPDSQANPSAAGAASGEKKNSQSILAMPQSGS 898
Db 57 NLAKEFHDEILSLKNVGERVYLEVEYELP----- 86
Qy 899 PEPESIRNTRSRSTPAIFASDPATCPILPGCEETIEISKGRG--LGSLVGGSD--TL 953
Db 87 --PVSLQSS-----VMF-----RTVEVTLHKGNTPGFVIRGGAHDDRK 125
Qy 954 LGAFITHEVEYERGAACKDRGLWAGQILEVNGIDLRKATHDEAIVLROTPQVRVRLT-Y 1012
Db 126 SRPVVITCVPRGGPDDRECTIKRPGDRLLSVDIRLLGTTTHAEMSLKQCGEATLIEY 185
Qy 1013 RDEAPKEEVCTLTIELOKKRGKLGISYVK--RNDTGVSIVKGIADPRGRIL 1070
Db 186 DVSAMDVAATASCPRLVEVAKTPGASLGVALTTSVCCNQVYIDIKSASIAIDRCGLH 245
Qy 1071 QGDQIILVNGEDVRNASEAVALLCSLGTVLEV-----GRIK-AGPFI--- 1115
Db 246 VGDHILSIDGTSMEYCTLAETQFLGNTDQVLELPHHQTLALKGPDPHVIKIRSQRQ 305
Qy 1116 -----SERRPSQISOVE 1128
Db 306 LPMDPWASSQSVHTNHNHPHDPCHVPALGFPKALTPNSPRAWSSSSPTSMSAVSL 365
Qy 1129 GSLSFTPLSGSSTS-----ESLESSKKNALASEIOGLRTVE--MKK 1170
Db 366 SSLNMGTLPRSLYSTSPRGTMRRRLKKDKFSSLSLASTVGLAQVHTETTEVYLTA 425
Qy 1171 GPTDSIGISYAGV--GSPLDGVPIFTAMMPTGVAOTOKLRVGRIVTIGSTEGMTH 1229
Db 426 DPTVGTGIOLOGSVFATELISPLISYIFADSPAERCGLVQIGDVMALNGIPFEDSTF 485
Qy 1230 TQAVNLKNAAGS-----TEMQVAGDVSVYTGHHQBPASSLSFTGLTSTSIFQDL 1283
Db 486 EEAQQLLRDSITSKYTLIEFDVAVESVIPSSTFHVKLKPKHSVEL-GITISSPSSRRP 544
Qy 1284 GPP-----OCK----- 1289
Db 545 GDLVVISDIKKSVAHRTGTLELDKLLAIDNIRLDSCEMADAVQILOQCEDEVKIKIRK 604
Qy 1290 -----SITLER--GPDGLGFSIVGSGPHGLPIYVTKVFAKGAASED 1331
Db 605 DEEDNSDEQESSGAIITYVELKRYGGLGITIS-----GTEPPDPPIISSLTNGLAERT 659
Qy 1332 GRLKRGDQIIVANGOSLEGVTHEEAVAILKRTKGYTLMV 1371
Db 660 GAIHGDRILAINSSSLKGRPLSEDIHLLQMAGETVTLKI 699

RESULT 9
US-09-045-632-16
; Sequence 16, Application US/09045632
; Patent No. 6001575

```

GENERAL INFORMATION:
APPLICANT: Huganir, Richard L.
TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,632
FILING DATE: 19-MAR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,016
FILING DATE: 19-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Corless, Peter F.
REGISTRATION NUMBER: 33,860
REFERENCE/DOCKET NUMBER: 48147/1699-CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1018 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-045-632-16

Query Match 4.88; Score 335; DB 3; Length 1018;
Best Local Similarity 22.55; Pred. No. 1.4e-18;
Matches 171; Conservative 90; Mismatches 261; Indels 238; Gaps 22;

787 LELPDDG-GGIAISEEDTLGVY-----IKSLFHEGVAATDGRKVGDDLLAVDE 838
3 VELAKKEETIGC-----TVSGIDKDKGRVSNLROGGIAANSDDLDVGDYIKAVNGI 56
839 IVGVPIEKFISLKTAKMTVKLTIHENPDQAVPSPAAGASEKKNSSQSLWVPOGS 898
57 NLAKFRHDEISILKKNVGERVYLEVEYELP----- 86
899 PEPSIRTSNASSPAPFASPATCPPIIPGCEITIEISKGTG--LGISTVGGSD--TL 953
87 --PVSIGSS-----VWF-----RTVEVTLHKEGVTGFEVINGAHDDRKNK 125
954 LGATIHIEVEGAACKGRGLMAGDQILEVNGIDLRKATHEAIVNLRQTPORVRLT-Y 1012
126 SRPVYITCVRRGGPDGRGTTIKPDRLSLVUGIRLLGTTHEANSLIKQCCQERTLLLEY 185
1013 RDEAPYKEEVEDPRLTTELQKPKGKGLSLVGR--RNDTGVFVSDIYKGGIADPDGRLI 1070
186 DVSAMDVAATASGPLVLEVAATPGASLGVALTSCCKKQIVYIDKIKSASIAIRGALH 245
1071 QGDDILLVNGEDVARNASOEAAVALKCSLGVTLLEV-----GRIK-AGPRH----- 1115
246 VGDHILSIDGTSMEYCTLAETQFLAGTTQVLEILPHHQTRIALKGPDPHVKIORSDRQ 305
1116 -----SERPSQTSQVSE 1128
306 LRPDPMASSQCCVHTNHHNPHDPHCHVPALGPALKALTPNSPPAMVSSSSPTSMASVSL 365

1129 GSLSFTPELSSSTS-----ESLESKKNALASEIGLRTVE--MKK 1170
366 SLSMNGTIPRLSLYSPGTMRRKRRKKDKSSLSIASSTVGLAGOVHETHEVILTA 425
1171 GPDSLGSISAGV-GSPGLGVPIFIAMHPTVAAQOKLRVDRIVTIGTSTEGMTA 1229
426 DPVTCFGILOGSVFATETLSSPLISYIEADSPAEKRGVLIQIDRWAIINGIPTEDSTF 485
1230 TQAVNLKNAGS-----IEMOVYAGDVSVYGHODEPSSLSLSTGLSTSTIFDDDL 1283
486 EEAQQLNDSSITSKYTLEIFDVAESVIPSSGFVHKLPRKHSVEL-GITISSPSSRRP 544
1284 GPP-----OCK----- 1289
545 GDPVLVSDIKKGSVAHRTGLELQDKLAIIDNIRLDCSMEDAVQIIQCCEDLYKRIK 604
1290 -----SITLER--GPDGLGFSIVGVGSPHGDLPYKTVFANGAASED 1331
605 DEDNSDEQESSGAIIVYELKRYGPGIGTIS-----GTEEPFDPIITISLTKGLAERT 659
1332 GLRKGDIIVNGOSLEGVTHEEVAVALKRTKGTVTLMV 1371
660 GAHIGDRILAINSSSLKGPLSEDHLLQWAGETVTLKI 699

RESULT 10
US-09-045-632-32
Sequence 32, Application US/09045632
Patent No. 6001575
GENERAL INFORMATION:
APPLICANT: Huganir, Richard L.
TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,632
FILING DATE: 19-MAR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,016
FILING DATE: 19-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Corless, Peter F.
REGISTRATION NUMBER: 33,860
REFERENCE/DOCKET NUMBER: 48147/1699-CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 1061 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-045-632-32

Query Match 4.88; Score 335; DB 3; Length 1061;

Best Local Similarity 22.5%; Pred. No. 1.5e-18;
Matches 171; Conservative 90; Mismatches 261; Indels 238; Gaps 22;

```

OY 787 LELPKDQG-GIGAISEEDTSLGVI-----IKSLTEHVAATDGRKLVGQDILAVDE 838
DB 3 VELMKKEGTTLC-----TVSGIDKDGKPRVSNLRQGIARSDQDLVDGYIKAVNGI 56
OY 839 IYVGYPIEFILKTAKTIVLTTHAENPDQAVPSAAGAAGEKNSQSILMPQSS 898
DB 57 NIAKFRHDEIISLKNVGRVVELEYELP----- 86
OY 899 PEPEIRNTSRSTPAIFASDPATCPRIIGCEETIEISKRTG--LGLSIVGSD---TL 953
DB 87 --PVSLQSS-----VMF-----RIYEVTLHKEGNTGFGFIRGAHDDRK 125
OY 954 ICAFIITHEVEEGAACKDGRMLNAGQILEVNGIDLRKATHEAIVNLQTPORVRLTL-Y 1012
DB 126 SRPVYITCVRPGPDRECTIKPRDRLLSVDEIRILGTTNHAEMSLIKOCGCEATLLIEY 185
OY 1013 RDEAPYKEEVCDDTLIELQKKRPGKGLSIVGK--RNDTGVFVSDIVAGIADPDGRLI 1070
DB 186 DVSANDSVATAGPLVEAKTPGASLGVALLTVCCNKOYVIDKIKASIAIDRCALH 245
OY 1071 QGDQILLVNGEDVRNASEVAVALKCSLGVTVLEV-----GRIK-AGPEH----- 1115
DB 246 VGDHLISIDGTSMEXCTLAETQFLGNTTDQKLEILPHNOTRLAKGPDHVKIQRSRQ 305
OY 1116 -----SERRPSQTSQVSE 1128
DB 306 LPWDPMWASSQCVHTNNHNNPHPHDCHRVAPALGPRAKLRPNSPRAVSSSPTSMSAYSL 365
OY 1129 GSLSFTPLSGSSTS-----ESLESSKKNNALASEIGLRTVE--MKK 1170
DB 366 SSLNMGTLPRSLXSTSPRGTMMRRLLKKKDFSSLSLASTYGLAQVYHTEETEVYLA 425
OY 1171 GPTDLSIGSIAGV--GSPGLDVPFIAMMPTGVAQOTKLRVGRIVTICSTEGMTN 1229
DB 426 DPTVGTGIGQDSVFATERLSSPRLSYIEADSPARCEVQLQIDRVMAINGIPEDSTF 485
OY 1230 TQAVNLKNAAGS-----IEMQYVAGGVSVYTGHHQBPASSLSFTGLTSTISITQDL 1283
DB 486 EKANOLLRDSITSKYVTLIEFDVASEVIPSQGFHVKLPRKHSYBL-GITISSPSSRP 544
OY 1284 GPP-----OCK----- 1289
DB 545 GDEPLVSDIKKSVAHRTGTLELQKLLIDNIRLDSCEMEDAVQILOCEDELVLKIKR 604
OY 1290 -----SITLER--GPDGLGFSTVGGYSGPHGDLPIYVKTVEFAKGAASED 1331
DB 605 DEDNSEDESSGAILITVELKRYGPGIGTIS-----GTEEPDPITLISLRTGCLAERT 659
OY 1332 GRLKRGDQIIAANGOSLEGVTHEEVAAILKRTKGVTLAV 1371
DB 660 GAHIGDRILAINSSSLKGPLSEDIHLQMGFTVTLKI 699

```

RESULT 11
US-08-410-804-1
Sequence 1, Application US/08410804
Patent No. 5632994

GENERAL INFORMATION:

APPLICANT: Reed, John C.
APPLICANT: Sato, Takaaki
TITLE OF INVENTION: FAS ASSOCIATED PROTEINS
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cathryn Campbell
STREET: 4370 La Jolla Village Drive, Ste 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/410.804
FILING DATE: 27-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/259,514
FILING DATE: 14-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn
REGISTRATION NUMBER: P-LJ 1389
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
US-08-410-804-1

Query Match 4.8%; Score 334.5; DB 1; Length 610;
Best Local Similarity 23.0%; Pred. No. 6.3e-19;
Matches 155; Conservative 72; Mismatches 221; Indels 225; Gaps 21;

OY 842 GYPIKEFISLTKAKTIVLTTHAENPDQAVPSAAGAAGEK-----NSSQSL 891
DB 2 GSPSPVIS-----KATEETITDSNQSTKPRGISDVTYDSDRGSDMDKATYSSQDH 56
OY 892 WYPOSQPEPEIRNTSRSTPAIFASDPATCPRIIGCEETIEISKRTGLSIV---- 947
DB 57 QPQPKQ---ESSSVNTSNMKNFTFSSPPK---PGDIFVELAKNDMSLGSIVTVLSD 109
OY 948 -GGSDDL--GAFIITHEVEEGAACKDGRMLNAGQDILEVNGIDLRKATHEAIVNLQTP 1004
DB 110 KGVNTSVRHGCIYVAVIPQGAESDGRIRHKGDRVLAANGVSLGATHKQAVETLRNG 169
OY 1005 QVRRLTVLDEAPYKEEV---C-----DRLTI 1029
DB 170 QVHLLERQSGSTSKENHVPPTQCTLSDQNAOGPEVKKTTQYKDYSEVTEETFEV 229
OY 1030 ELQKKPKGLGLS-----IYKRNQDGVFVSDIVKGIADPDGRLIOGQDILLVNGED 1082
DB 230 KLEFN--SSGLGFSFREDNLIRQINASTIVRYKKLFPQGPAAESKIDVGDVILKNGAS 288
OY 1083 VRNASQEAVALKCSLGVTVLEVGRIRKAGP-----HSERRPSQT 1123
DB 289 LKGLSQOEIVSALRGPAVEFLLCRPPGVLPETIDTALLPLQSPAVLPNNSKDSQP 348
OY 1124 SQVSEGLSSTFTPLSGSSTSESLESSKKNALASIQGLRTVEAKKGTDSIGISIAG 1183
DB 349 SCVEQ-----STSSDENEMSDSKKQC-----KSPSRDSYSDSSG 384
OY 1184 VSPPLGDVPFIAMMPTGVAQOTKLRVGRIVTICSTEGMTHTQAVNLKNAAGSI 1243
DB 385 SGE-----DDLVT-----APAINSNSTWS 404
OY 1244 EMQYVAGGVSVYTGHHQBPASSLSFTGL-----TSTSIPODDLCP--- 1285
DB 405 ALHQTLSNNVSOASHHEAPKQOEDTICTMFTYQKIPMKPEREDSNPPLPPDMAFGS 464
OY 1286 --PQCKS-----ITLNGPQC 1299
DB 465 YOPQSESSASSMDKYHHIISEPTROENMTPLKNDLENHLDPELEVLLTTLKSEKG 524
OY 1300 -LGFSIVGGYSGPHGDLPIYVKTVEFAKGAASEDGRKRGDQIIAANGOSLEGVTHEEVA 1358

```

Db 525 SLGFTVTKG----NORIGCYHVDYI-QDPAKSDRLKPGDRILKIVNDTDYNTMHTDAVN 579

QY 1359 ILKRTKGTVMV 1371

Db 580 LLRAASKTVRLVI 592

RESULT 12

US-08-259-514-1

Sequence 1, Application US/08259514

Patent No. 5747245

GENERAL INFORMATION:

APPLICANT: Reed, John C.

APPLICANT: Sato, Takaaki

TITLE OF INVENTION: FAS ASSOCIATED PROTEINS

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Cathryn Campbell

STREET: 4370 La Jolla Village Drive, Ste 700

CITY: San Diego

STATE: California

COUNTRY: United States

ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/259,514

FILING DATE: 14-JUN-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn

REGISTRATION NUMBER: 31,815

REFERENCE/DOCKET NUMBER: P-LJ 9954

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 535-9001

TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 610 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-259-514-1

Query Match 4.88; Score 334.5; DB 1; Length 610;

Best Local Similarity 23.08; Pred. No. 6.3e-19;

Matches 155; Conservative 72; Mismatches 221; Indels 225; Gaps 21;

QY 842 GYPIKEFISLTKAKMYVKTIIHAENPDQAVPSAAGASGEK-----NSSQSL 891

Db 2 GSPSPSVIS-----KATKEKFTDSNOSKTKKPGISDVTDSGDSDMDATYSSSDH 56

QY 892 MWPGSGSPESINNTSSSPAFASDPATCPIIPGCEITIEISKGTGIGLSIV---- 947

Db 57 QTPKQ-----ESSSVNTSKNMFKTFSSSPK-----PGDIFEVELAKNDNSLGISVTLFD 109

QY 948 -GGSDDL--GAFIIHEYEGACAKGRLMAGQILEVNGIDLRKATHDAIVLQTP 1004

Db 110 KGVNTSVRHGIIYKAVIPQGAESDGRIRKGRVLAIVNVSLEGATHKQAVETLRNTG 169

QY 1005 QRVRLTIYRDAPYKEEV-----C-----DLTI 1029

Db 170 QVALLLEKGGSPSKHEVPTPOCTISDONAGQGPEKAKKTVQVYDVSFVTEENFEV 229

QY 1030 ELAKKPKGLGLS-----IVGRNDYGVFSDIVKGIADPDGRLIOGDIILVNGED 1082

Db 230 KLFKN-SSGIGFSFSRBDNLIPEQINASIYVKKLPFGPAAESGKIDVGVILKVGAS 288

QY 1083 VRNASEVAALLKCSIGTIVLEVGRIKAGP-----HSERRPSQT 1123

Db 289 LKGLSQQEVIALKRGTAPEVFLLCRPPGVLPEDIRALLPLQSPAQVLPNSSKDSQP 348

QY 1124 SQVSEGLSFTPLPSSGSSSESLESSKKNALASEIOGLFTVEKKGPDIISLGIAGC 1183

Db 349 SCVEQ-----STSDENEMSDSKKQC-----KSPSRDSYSPSSG 384

QY 1184 VQSPGLGVPIFIAMHPTGVAAGQOKLAVGDRIVTIGSTEGMTHQAVALLKNASGSI 1243

Db 385 SGE-----DDLVY-----APANISNTWSS 404

QY 1244 EMQVAVAGDVSVYTHHQEPASSLSFTGL-----TSTSIPODLGP--- 1285

Db 405 ALHQTLNMTVSAQASHHAPKQSDITCTMFTYPOKIPNKPFEEDSNPSPLPDMARGOS 464

QY 1286 --POCKS-----ITLERPGD 1299

Db 465 YQPOSESASSSSMDKIHIIHISEPTROENWTPLKNDLENHLEDELEVEILLITLIKSG 524

QY 1300 -LGFSLVGGYSPHGDLPYKYTFAPKGAASEDRKRGDQIIAVNGQSLGVTHEEAVA 1358

Db 525 SLGFTVTKG----NORIGCYHVDYI-QDPAKSDRLKPGDRILKIVNDTDYNTMHTDAVN 579

QY 1359 ILKRTKGTVMV 1371

Db 580 LLRAASKTVRLVI 592

RESULT 13

US-08-858-311-1

Sequence 1, Application US/08858311

Patent No. 5876939

GENERAL INFORMATION:

APPLICANT: Reed, John C.

APPLICANT: Sato, Takaaki

TITLE OF INVENTION: FAS ASSOCIATED PROTEINS

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Cathryn Campbell

STREET: 4370 La Jolla Village Drive, Ste 700

CITY: San Diego

STATE: California

COUNTRY: United States

ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/858,311

FILING DATE: 27-MAR-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn

REGISTRATION NUMBER: 31,815

REFERENCE/DOCKET NUMBER: P-LJ 1389

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 535-9001

TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 610 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-858-311-1

Query Match 4.8%; Score 334.5; DB 2; Length 610;
 Best Local Similarity 23.0%; Pred. No. 6.3e-19;
 Matches 155; Conservative 72; Mismatches 221; Indels 225; Gaps 21;

QY 842 GYPIKFTSLKTAQMYKLTIAHNPDSQAAPSAAGASGK-----NSSQL 891
 DB 2 GSPSPVLS-----KATEKFTDSNOSKTKKPGISDVTYSDRGSDMDKATYSSQDH 56
 QY 892 MYPOGSEPEIRNTSSSTPAITAPDPAICPIIPGCTTETIKSGRTGLSTV----- 947
 DB 57 QTPKO---ESSSVNTSKMKNFKTSSSPK---PGIIFVELAKNDNSLISVTFED 109
 QY 948 -GGSFTLL-GAFITHEVEEGACDGLMAGDQILEVNGIDLKRTADEAINVLKOTP 1004
 DB 110 KGVNTSVRHGIIYKAVIPQGAESDGRINHGDVLAIVGSLGATHKQAVELRNTG 169
 QY 1005 QVRLLTYRDEAPYKEEV---C-----DRLTI 1029
 DB 170 QVHLLLEKGSPTSKENHVPYPOCTLRHONAGOGPEKKTQVKDYFVTEENTFEV 229
 QY 1030 ELQKRPKGLGIS-----YGRNDTGVFVSDIVKGIADPDRLIOGDIILVNGED 1082
 DB 230 KLFKN-SSGLFSPSREDNLPEQINASIVRYKLLFPGPAESGKIDGVILKVNQAS 288
 QY 1083 VNNASOEVAALLKCSLGTVLEVRKAGPE-----HSERRPQOT 1123
 DB 289 LKGLSQOEYISALRGTAPRVLLLCRPPPGVLPETIDTALLTLPQSPAOVLPNSSDSSOP 348
 QY 1124 SOVSEGSLSFTPLSGSTSESSSKKNALASEIOGLRTVENKKGPTDSLIGSIAGG 1183
 DB 349 SCVEQ-----STSDENEMSDKSKQC-----KPSRRDSYSDSG 384
 QY 1184 VGSPLGDPVPIFAMMPPGVAAOTOKLRYGRIVITICSTEGMHTQAVNLKASGSI 1243
 DB 385 SE-----DDLVT-----APANISSTWS 404
 QY 1244 EMQVAVAGDVSVYTGHHOEPASSLSFTGL-----TSTSPFDDIGP--- 1285
 DB 405 ALHQLLSMNVSOAOSHHEAPKQEDTICMTFYPOKIPKPEFEDSNPSPLPPDAPGOS 464
 QY 1286 --POCKS-----TLERGPDG 1299
 DB 465 YOPGESASSSSMDKYHHIHSPTROENWPLKNDLENHLEDFEVELLTLTKSEK 524
 QY 1300 -LGEFIVGCGPHDPLRYKTVFAKGAASDGRKRDQIIAIVNGSLEBVTHEEAVA 1358
 DB 535 SLGFVTYKQ---NORIGCYHDI-ODPAKSDGRKPDRLIKVNDTDTMKTHTDAVN 579
 QY 1359 ILKRTKGVTLNV 1371
 DB 580 LIRAAKTYRLVI 592

RESULT 14
 US-09-045-632-50
 Sequence 50, Application US/09045632
 Patent No. 6001575
 GENERAL INFORMATION:
 APPLICANT: Huanli, Richard L.
 APPLICANT: Dong, Hualing
 TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
 TITLE OF INVENTION: GRIP-RELATED MOLECULES
 NUMBER OF SEQUENCES: 105
 CORRESPONDENCE ADDRESS:
 ADDRESSSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
 STREET: 130 Water Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/045,632
 FILING DATE: 19-MAR-1998
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/041,016
 FILING DATE: 19-MAR-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Corless, Peter F.
 REGISTRATION NUMBER: 33,860
 REFERENCE/DOCKET NUMBER: 48147/1699-CIP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-523-3400
 TELEFAX: 617-523-6440
 INFORMATION FOR SEQ ID NO: 50:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1050 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-045-632-50

Query Match 4.8%; Score 334.5; DB 3; Length 1050;
 Best Local Similarity 21.9%; Pred. No. 1.6e-18;
 Matches 203; Conservative 129; Mismatches 321; Indels 275; Gaps 42;

QY 525 VELMREPKSIGISIVGR-----GNGSR---LSNGEVR---GIFIKVL 564
 DB 89 VELIRBESSTGLTSSGTDKGRVSNLPRGGLAASDLNVGDIYRSNGILFRLR 148
 QY 565 EDSPPAKNGTL--KPDRIIVEAPSOSESEPERKAPLCVPPPPSAFAMGSDHTQSSASK 622
 DB 149 HDEIL---TLKNKGENV---LEVEYE---LPPAP-----ENNRI 182
 QY 623 ISQVDKEDERCYSKNIIRERYGTLTGLHMI TELEKHSGLSLAG---NKDRSRNV 678
 DB 183 ISKTYD-----VSLKBEKNSFGFVLRGCAHEDLHRSRLV 217
 QY 679 FIVGIDPAGAKGDRLOADELEINGOLLYGRSHON--SIIKCAPSKVYIIFRKND 737
 DB 218 -LTYRPGPARREGSLVGDRLSTIXGIPLGASHAIAIATLOQ-----SHE 265
 QY 738 AVNQAVCPGNAVEPLPSNSENLOKTEPTVTSDAAVDLSSFNVOHLELPKDOGGLG 797
 DB 266 ALFYV-----EYDVPDPVAMNSGPLVVEIAKTPGSALGIS 302
 QY 798 IAISEDTLSGYIISLIEHGYAATDGRKVGDOILAVDELIVGYPLEKFTSLKTKM 857
 DB 303 LTTGSHRNKPATITDIRIPASVVDKNGALHGEHTIADTSTECSSVEATKLASVTE 362
 QY 858 TYKLTIAHENPDSO---AVPSAAGAAGEKKNSQSLWVPOGSEPEPSIR----- 905
 DB 363 KVRLEI-LPAPOSRPLKPEEAVRIORSEQLHMDP-CVPSCHSPRSHCAPTAPGGO 420
 QY 906 --NTRSSTPAIFASDPAICPIIPGCTTETISKG---RT----- 940
 DB 421 DQSRVSSSTP---FSSPMNPAFP-CANASTLPROMSPRTTAGRRORRKHRSLSLA 476
 QY 941 -----GGLSTVGG---SDTLGAFITHEVEEBAACKD 971
 DB 477 SSTVPGQIVHFTTEVYVLCGDPISGFLQIGAFATETLSSPLVFTEPDSAPARC 536
 QY 972 GRLMAGDOILEVNGIDLKRTADEAINVLKOT--PQVRLTYRDEAPYKEEV---CDT 1026
 DB 537 GILQVGDVLAINGIATEDGTMEANQILRDAALARKIVLEIEFPVA---ESVIPSST 592
 QY 1027 LTELQKRPKGLGSI---VGRNDTGVFVSDIVKGIADPDGRLIOGDIILVNGEDVR 1084

```

Db 593 FHVKLPKRGVGLGTTISSASRRKGEPLISDIKKSVAHRTGTLPEQDKLLAIDNIRLD 652
QY 1085 NASOE-AVAALLKCSLGTVLEVGRIKAGPHSERPSQTS--OVSEGLSSTFFPLSGS 1141
Db 653 HCPMEYAVQILPOCE-DLVKTKI-----RKEDNSDEQESSGAVS-----691
QY 1142 STSELESSSKKNMLASBIOGLRTVEKK--GPTDSIGISIAGVGSPLGDVPFIAMH 1199
Db 692 -----YVELKRYGCP--LGITIS-GTEEPFD--PIIISGLT 723
QY 1200 PTGYAOTOKRLVGDRIYITIGTSTEGTHTQAVNLKNAGSIEMQV-----A 1249
Db 724 KRGLAERTGASMLGTAWPXSYSVLKGRPLSEALHLLQVAGEYTLKIKQDLRPLRQ 783
QY 1250 GGDVSVYTGHHQEPASSLSFTGLTSTISFODDLGPPQCKSITLERGPDGLGFSIV-GGY 1308
Db 784 SGSLSEASDVDEDPPEALKG--GLTTHF-----SPAVPSV--DSAVESGSSATEGCF 833
QY 1309 GSPHGDLPYKTYFAKAGASEDGR LKR 1336
Db 834 GSGS-----YTPQVAVRSVTPQEWRRSR 857

RESULT 15
US-09-045-632-49
: Sequence 49, Application US/09045632
: Patent No. 6001575
: GENERAL INFORMATION:
: APPLICANT: Huganir, Richard L.
: APPLICANT: Dong, Hualing
: TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
: NUMBER OF INVENTION: GRIP-RELATED MOLECULES
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
: STREET: 130 Water Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/045,632
: FILING DATE: 19-MAR-1998
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/041,116
: FILING DATE: 19-MAR-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Corless, Peter F.
: REGISTRATION NUMBER: 33,860
: REFERENCE/DOCKET NUMBER: 48147/1699-CIP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-523-3400
: TELEFAX: 617-523-6440
: INFORMATION FOR SEQ ID NO: 49:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1050 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-045-632-49

```

Query Match 4.7%; Score 331.5; DB 3; Length 1050;

Best Local Similarity 21.8%; Pred. No. 2.9e-18;

Matches 202; Conservative 130; Mismatches 321; Indels 275; Gaps 42;

QY 525 VELKREPSKLSIGISIVGR-----GMGSR---LSNGEYMR---GIFIKHVL 564

```

Db 89 VELIKREGSTIGLITSGTQDKGKPRVSNLRPGGLAARSDLLNAGDYIRSVNGIRLRLR 148
QY 565 EDSFAGKNGTL--KPGRIYAPASQSESEPEKALCSVPPPPAFAMGSDHQSASK 622
Db 149 HDEIT--LLKNVGERVY-----LEVEYE-----LPPAP-----ENNRI 182
QY 623 ISODPKEDFEFGYSWKNIIRERYGTLTGELHMIIELEKSHGSLAG-----NKDRSMV 678
Db 183 ISKTVD-----VSLYKENSNGFPLRGAMHDLKSRPLV 217
QY 679 FIVGIDPNAAGDKRLQIADLELINOQIILYGRSHQIA--SIIKCAPSKYKIIIFIRNKD 737
Db 218 -LTYVRPGGPANREGSLKVGDRLLSIXGIPLHGASHATATLQOC-----SHE 265
QY 738 AVNOMAVCPGNAVEPLPSNSENLOKETEPEVTTSDAVDSFRKNOHELPRKDGGLG 797
Db 266 ALFOY-----EIVYATPDYANASGLVVEIAKTPGSALGIS 302
QY 798 IAISEEDTLGCVIIKSLTEHGAATDGR LKVGDOI LNAVDEIVVGYPIEKELSLKTAKM 857
Db 303 LTTGSHRKRPAITIDRIKIPASVVDKNGALHAGEHILAIIDGTSTEHCSLVEATKLLASYTE 362
QY 858 TVKLTIRAHENDSQ--AVPSAAGASGEKNSOSLAVPQSGSPESIR-----905
Db 363 KVRLEI-LPAQOSRRPLKPEPAVRIRQSEQLHMDP-CVPSCHSPRSHCAPTWAPGQ 420
QY 906 --NTSRSTPAIFADPATCPIIPGCEITIEISK--RT-----940
Db 421 DOSRSVSTP--FSSPIMNAPF-CANASTLPFGPMSPTTARORRORREHRSLSLA 476
QY 941 -----GLGSLVGG--SDTLGAFITHEVEEGAACKD 971
Db 477 SSTVPGQGIYHTEFEVVLGCDPLSGGLQLOGIFATEILSSPPLVRFTEPSPAPERC 536
QY 972 GRMAGQIILEVNGIDLRKATHDEAINVLRQT--PQVRRLTYNDEAPYKEEV---CDF 1026
Db 537 GLLOVGDRLVAINGIATEDGTEQANOLLRDALARKIVLEIEFDVA---ESVIYPSGT 592
QY 1027 LTILOKPKGKGLSLI--VGRNDYGVFVSDIYKGIADPDGR LIOGDOI LLVNGEDVR 1084
Db 593 FHVKLPKRGVGLGTTISSASRRKGEPLISDIKKSVAHRTGTLPEQDKLLAIDNIRLD 652
QY 1085 NASOE-AVAALLKCSLGTVLEVGRIKAGPHSERPSQTS--OVSEGLSSTFFPLSGS 1141
Db 653 HCPMEYAVQILPOCE-DLVKTKI-----RKEDNSDEQESSGAVS-----691
QY 1142 STSELESSSKKNMLASBIOGLRTVEKK--GPTDSIGISIAGVGSPLGDVPFIAMH 1199
Db 692 -----YVELKRYGCP--LGITIS-GTEEPFD--PIIISGLT 723
QY 1200 PTGYAOTOKRLVGDRIYITIGTSTEGTHTQAVNLKNAGSIEMQV-----A 1249
Db 724 KRGLAERTGASMLGTAWPXSYSVLKGRPLSEALHLLQVAGEYTLKIKQDLRPLRQ 783
QY 1250 GGDVSVYTGHHQEPASSLSFTGLTSTISFODDLGPPQCKSITLERGPDGLGFSIV-GGY 1308
Db 784 SGSLSEASDVDEDPPEALKG--GLTTHF-----SPAVPSV--DSAVESGSSATEGCF 833
QY 1309 GSPHGDLPYKTYFAKAGASEDGR LKR 1336
Db 834 GSGS-----YTPQVAVRSVTPQEWRRSR 857

```

Search completed: July 12, 2001, 14:40:58
Job time: 201 sec

Fri Jul 13 15:00:07 2001

us-09-502-698-1.rai

Page 13

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2001, 14:40:26 ; Search time 41.92 Seconds

(without alignments)
1453.414 Million cell updates/sec

Title: US-09-502-698-2

Perfect score: 5085
Sequence: 1 MLOWNSKESERTINAKGN.....EEAVALLKRTKVTLMVLS 1005

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_0601.*
1: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5085	100.0	1005	20	AAV04731
2	5085	100.0	1373	20	AAV04730
3	5085	100.0	2000	20	AAV04732
4	5035	99.0	2070	20	AAV04733
5	5018.5	98.7	2037	21	AAV53753
6	4080.5	80.2	856	21	AAV01383
7	3840	75.5	763	20	AAV04741
8	2133.5	42.0	1881	20	AAV24025
9	1501	29.5	318	20	AAV74150
10	1127	22.2	251	20	AAV04738
11	1089	21.4	272	20	AAV04739

12	1089	21.4	319	20	AAV04740
13	1089	21.4	1239	20	AAV04734
14	928	18.2	206	20	AAV74151
15	583	11.5	632	21	AAV6669
16	583	11.5	632	22	AAV87545
17	583	11.5	632	22	AAV85212
18	580	11.4	632	21	AAV01385
19	509	10.0	2485	21	AAV39343
20	505.5	9.9	674	20	AAV2953
21	500.5	9.8	2466	16	AAV71498
22	500.5	9.8	2466	19	AAV75999
23	500.5	9.8	2466	21	AAV90272
24	443.5	8.7	590	20	AAV04736
25	399	7.8	817	20	AAV48101
26	399	7.8	817	20	AAV30137
27	399	7.8	849	19	AAV48102
28	397	7.8	1161	21	AAV03832
29	396.5	7.8	1112	21	AAV03833
30	395	7.8	1277	21	AAV03149
31	385	7.6	767	21	AAV2138
32	376	7.4	344	22	AAV55834
33	376	7.4	344	22	AAV57623
34	376	7.4	344	22	AAV58035
35	372	7.3	882	18	AAV34662
36	359	7.1	1113	22	AAV4404
37	357	7.0	1112	19	AAV73061
38	347	6.8	744	21	AAV71410
39	339.5	6.7	396	18	AAV34666
40	339.5	6.7	1049	19	AAV73062
41	334.5	6.6	610	17	AAV94649
42	315.5	6.2	414	22	AAV36687
43	303.5	6.0	227	21	AAV53276
44	303.5	6.0	419	22	AAV64997
45	298	5.9	207	22	AAV57625

ALIGNMENTS

RESULT 1
ID AAV04731 standard; Protein: 1005 AA.
AC AAV04731;
DT 06-JUL-1999 (first entry)
DE Mature protein containing PDZ domain from clone 38-2-1.
XX PDZ domain: gene expression; human umbilical vascular endothelial cell;
XX HUVEC; stimulation; tumour necrosis factor; TNF; protein binding;
XX cell; proliferation disorder; cancer.
XX Homo sapiens.
XX WO9907846-A1.
XX 18-FEB-1999.
XX 12-AUG-1998; 98WO-JP03603.
XX 19-JUN-1998; 98JP-0189944.
XX 12-AUG-1997; 97JP-0230356.
XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX Funahashi S, Miyata S;
XX WPT. 1999-167423/14.
XX N-PSDB; AAX29908.
XX Protein containing PDZ domain, whose expression is enhanced by TNF
XX stimulation - plays an important role in protein/protein

PT interactions and is used for screening for proteins for use in
PT treatment of cell proliferation disorders such as cancer

PS Claim 1; Page 84-87; 240pp; Japanese

CC This sentence represents the mature portion of a new protein containing
CC a PDZ domain encoded by the clone 38-2-1, whose expression in human
CC umbilical vascular endothelial cells (HUVEC) is enhanced by stimulation
CC with tumour necrosis factor (TNF) alpha. The new protein is used to
CC identify proteins which bind to it (particularly to the PDZ domains) and
CC the genes encoding them, for use in the treatment of cell proliferation
CC disorders such as cancer.

SQ Sequence 1005 AA;

Query Match	100.0%	Score 5085	DB 20	Length 1005
Best Local Similarity	100.0%	Pred. NO	0	
Matches 1005	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Qy	1	MLQWVSKSEFPRTJNINAKGNSLSLQMTYSANKDQJGMIYRSTIHGGAISRQRIAGCII	60
Dd	1	mlqvwsksefertilnakgnslsmtysankdjqgmlyrstihggaisrqrriagcili	60
Qy	61	SINEBSTISVTNAQARAMLRHSLIGDPDKITVYPAHLEEFKISLQCSGRVMAIDIFS	120
Dd	61	sinebstisvtnaqaramlrhslrthsilgdpdkitcyphaehleefkislqcgsvrmaidifs	120
Qy	121	SYTRDRIPELPEREGEGSESELONTYVNMNNOBRVRLMEPERSKISGISYVGRGMSR	180
Dd	121	sytrdrilpelpereregeseeseloqtyvnmnnobrvrlmeperskislgisylvgrgmshr	180
Qy	181	LSNGEVMKGIETIKHVLHEDSPAGKNGKTAPCDRIYEAPOSSESEPEKAPLCSVPPPPSAF	240
Dd	181	lsngewmgzikihvledspagngkllkpgdriyeapsesepekaplcsvppppsaf	240
Qy	241	AEMKSDHTQSSAKISQDVKDEDFEGYSKMNIRRRYGTTLGELHMLEFGHSGSLGSLA	300
Dd	241	aemsgsdhtqssaasklsqdvkdedfegyswknirrrygttlgelhmlfeghshsgslgsls	300
Qy	301	GNKRSRMSVTVIGIDPAGGAKCGRROIIDELLEINGOLLYGSHONASSITIKCAPSKV	360
Dd	301	gnkdrsrmsvtviyidpagaagkgrriqiadelletnrgqlllygshbnassilikcapskv	360
Qy	361	KIIFIRKNDAVNQNAVCPGNAVEFLPNSNENLQNKETEPYVTSDDAVIDLSSEFKNOHLE	420
Dd	361	kififrnkdavnqnavcpogneveflpnsenlnqketepervtsdaavidlsfknvghle	420
Qy	421	LPKQOQGLGIAISEEDTISGYITIKSLTEHGYAATDGRKAYGDOGLAVNDELIVYGPPIEKF	480
Dd	421	lpkqdgglglaiseedtlsgyilksltehgyaaedgrlkaygdglavndellvvgypriekef	480
Qy	481	ISLTKTKMKMYKLTIIHAENPQSOAVPBAAGAAASECKNSNSSLVCPVSGSPEPESINTS	540
Dd	481	islltktkmkykltihenpdsqavpbaagaasecknssnsslvmpqsgspepesints	540
Qy	541	RSSTPAIFASDPATCPIIPGCEETIETLSKGTGIGLSIVGSDTLGCAFIHNEYEEGAA	600
Dd	541	rsttpaifasdpatcpiipgceetietlskgrtgiglsivgsdttlgaflihveyeegaa	600
Qy	601	CKDGRUWAGDOITLEVNGSIDLRKATHADAIVLAKOTPORVRLTLYRDAAPKKEEVCDTLT	660
Dd	601	ckdgrlwagdqitlevngsidlrkathadaivlqtpbrvrltlyrdeaprkkeevecdtlc	660
Qy	661	IELQKKKGKGGIGLSIVCKRNDTGVFVSDIYKGGIADBDGRILGODOILLANGEDVRNASQ	720
Dd	661	ielqkkrgkgiglsivckrndtgvfvsvdiykggiaadbgrilgodoillangedvrvnasq	720
Qy	721	EAVVAALFKCSLGTYTTELEGRITAKAPFHSERRPQSQTQVSGSLSSTFFPLSGSSTSELE	780
Dd	721	eavvaallkcslgtyttelegritakapfhserrpqsqtqvsgsllsstffplsgsstsele	780
Qy	781	SSSKKNMALASEIGLRTVEEMKKGFTDLSGISIAGGVSPGLADVPITAMINPTGVAAQTQ	840

Db	781	ssskknaaaseiglrivremkkgpctdsjgsiaagvsygpjdrpifiammhprgvaatq	840
Qy	841	klrvgrdriwttcgtstegmthtqaavnlknknsigsiekvvaagdvsvtghoeppassl	900
Db	841	klrygddivlvcgstegmctgvnlhlnaasglcmqvaagdvsvtlbhepassl	900
Qy	901	sftglststifodolcgpockstilerpglgsiyvgsgspgldpripvktvfaakaa	960
Db	901	stfgltsstisqddlppgcstilerpglgsiyvgyspdpripvktvfakaa	960
Qy	961	sedrklrkgdiiivnqsgslsgytheaavnlkrtktyvlmvl	1005
Db	961	sedrklrkgdiiavngqslsgytheaavnlkrtktyvlmvl	1005

RESULT 2

ID AAY04730 standard; Protein; 1373 AA.

AC AAY04730;

DT 06-JUL-1999 (first entry)

Protein containing PDZ domain from clone 38-2-1.

KW PDZ domain; gene expression; human umbilical vascular endothelial cell; nitric oxide stimulation; tumor necrosis factor- α ; TNF- α ; protein binding

KW cell; prolife
xy

OS	Homo sapiens
XY	

PN WO9907846-A1
XX

PD 18-FEB-1999.
XY

PF 12-AUG-1998; 98WO-JP03603.
XX

PR	12-AUG-1997:	97JP-0230356
PR	19-JUN-1998;	98JP-0189944
PR		

PA (CHUG-) CHUGAT RES INST MOLECULAR MEDICINE INC
XX

Funahashi S., Miyata S.:
PI

XX
DR
WPT: 1999-167423/14

XX N-PSDB; AAX29908
DK
XX

PT stimulation - plays an important role in protein/protein

PT treatment of cell proliferation disorders such as cancer

aa
PS Claim 1; Page 77-84; 240pp; Japanese

CC This sequence represents

CC by the clone 38-2-1, whose expression in human umbilical vascular
CC endothelial cells (HUVEC) is enhanced by stimulation with tumour necrosis
CC factor (TNF) alpha. The new protein is used to identify proteins which
CC bind to it (particularly to the p12 domains) and the genes encoding them,
CC for use in the treatment of cell proliferation disorders such as cancer.
SQ
XX Sequence 1373 AA;

Query Match	100.0%	Score 5085	DB 20	Length 1373
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1005	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	1	M	A	N	K	E	S	E	R	T	I	N	A	K	N	S	I	C	M	T	V	S	A	N	K	R	G	L	M	I	V	S	I	T	H	G	A	I	R	D	R	I	A	I	G	C	T	L	60
Db	369	m	i	n	a	n	k	e	s	e	r	t	i	n	a	k	n	s	i	g	m	v	s	a	n	k	g	i	m	i	v	s	i	t	h	g	a	i	s	d	r	i	a	i	g	c	t	l	428

```

QY 61 SINEESTISTVNAQARAMLRRHSLIGPDITKITYVPAHLEEFKISIGQSGRMALDIFS 120
   |||||||
Db 429 slineestistvnaqaramlrrhsligpdikityvpaehleefkisygqsgymaldifs 488
QY 121 SYTGNDIPELPEREGEGESELQNTAYSNMNPARRVELMREPSKLSIGISVGRMGSR 180
   |||||||
Db 489 sytgndipelperegeegeseelqntaysnmnprrvelwrepsklsigisvgrmgstr 548
QY 181 LSNGEVWRCGFIKRVHVEDSPAGKNGTLKPGDRIVEAPRSQSESPERAPLCSVPPPPSAR 240
   |||||||
Db 549 lsngevwmrgfikhvledspagkngtlkpgdriveaprsqsesepkeraplcsvppppsar 608
QY 241 AEMGSDHTOSASAKISQDVYKEDDFGYSWKNIERRGTTLGELHMIIELEKGGSLGSLA 300
   |||||||
Db 609 aemgsdhtosasaakisdqdvkedefgyswkniergytlgelhmiielekghsglsls 668
QY 301 GNRKDRMSVFIYIDPNGAAGKDGRLQIADLELEINGQILYGRSHQNASIITKCAPSKV 360
   |||||||
Db 669 gnrkdrmsvfiyidpngaagkdgrlqiadelleingqillygrshqnasiltkcapskv 728
QY 361 KIIFIRNKDAVNQMAVCPGNAVEPLPSNSENLONKETEPTVTTSDAAYDLSSFRKNVQHE 420
   |||||||
Db 729 kifirnkdaavnqmavcpgनावेप्लप्सनेनलनकेप्टवत्सदावदलसफ्रकनव्हे 788
QY 421 LPKDGGGLGIAISEEDPTLSGVIIKSLTEHGVATDGRKLVGDOILAVDDEITVGYPIEK 480
   |||||||
Db 789 lpkdggglgiaiseedptlsvglksltehgvaaatdgrlkvgdqllavddeivvygplekf 848
QY 481 ISLLTKAKMTVKLTTHAENPDQAVPSAAGASGEKNSQSILMVPQSGSEPEESIRNTS 540
   |||||||
Db 849 islltkakmtvklthhaenpdsgavpsaagaasekknssqslmvpqsgsepesirnts 908
QY 541 RSTTRPAIFASDPATCPRIIPCCERTTIESKRTGLGSIYVGSOTLGAFTIHEVEEGAA 600
   |||||||
Db 909 rsttrpaifasdpactprilppccerttiegkgrtglsiyvgsdcllgaftihveegaa 968
QY 601 CKRGRLMAGQOILEVNGIDIRKATHDEAINVLTROTPORVLTLYRDPAPKEEVCOTLT 660
   |||||||
Db 969 ckrgrlmagqoilevngidirkathdeainvltrotprvltlyrdeaprykeevcotlt 1028
QY 661 IELQKRPKGLGSIYVGRNDTGVFVSDIYKGIADPDGRILQGDQILLVNGEDVNRASQ 720
   |||||||
Db 1029 ielqkrrpgkglisyvgrndtgvfvsdlvkgladpdrlllgdqlllvngedvnrnsq 1088
QY 721 EAVVALLKCSLGVTVLEVEGRIKAGPRHSEKRRPQTSQVSGSLSTFRPLSGSSTSESL 780
   |||||||
Db 1089 eavvaallkcslgvtvlevegrikagprfhsekrpqtsgvsgslsstfrplsgsstsele 1148
QY 781 SSSKKNALASEIGGLRTVEKKKGPRTDSIGTSIAGVSGSPDGDPRIETAMHPTGVAAQTQ 840
   |||||||
Db 1149 ssskknalaseigglrtvekkkgrptdsgtsiagvsgspdgdprietammhptgvaaqtq 1208
QY 841 KLVNGBRIVYICGSTEGMHTQAVNLKKNASGSIEMQVAVAGDVSVTGHQHPASSSL 900
   |||||||
Db 1209 klvngbrivlyicgstegmhtqavnllkknasgslmqvavagdvsvtghhqpasasl 1268
QY 901 SFTGGLTSTSTFQDDLPBPCKSTTLERBPGLGSPSYGVGSPHGDPIYVKTVPKAGAA 960
   |||||||
Db 1269 sftgglststfqqddlpbpcksttlerrbpdglsfsvgygspbhgdprlyvktvfkagaa 1328
QY 961 SEQGRILKRGDOITAVNGOSLEGVTHEEVAVALILKRTGTVTLAMVLS 1005
   |||||||
Db 1329 seqgrilkrqdgitavngqslgsvtheevaualilkrtkgtvltlmvls 1373

```

```

DE Protein containing PDZ domain from clone 38-2-1a.
XX
KW PDZ domain; gene expression; human umbilical vascular endothelial cell;
KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding;
KW cell; proliferation disorder; cancer.
XX
OS Homo sapiens.
XX
PN WO9907846-A1.
XX
PD 18-FEB-1999.
XX
PF 12-AUG-1998; 98WC-JP03603.
XX
PR 19-JUN-1998; 98JP-0189944.
XX
PR 12-AUG-1997; 97JP-0230356.
XX
PA (CHUGAI) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
PI Funahashi S, Miyata S;
XX
DR WPI; 1999-167423/14.
XX
DR N-PSDB; AAX29908.
XX
PT Protein containing PDZ domain, whose expression is enhanced by TNF
PT stimulation - plays an important role in protein/protein
PT interactions and is used for screening for proteins for use in
PT treatment of cell proliferation disorders such as cancer
XX
PS Claim 1; Page 158-167; 240pp; Japanese.
XX
CC This sequence represents a new protein containing a PDZ domain encoded
CC by the clone 38-2-1a, whose expression in human umbilical vascular
CC endothelial cells (HUVEC) is enhanced by stimulation with tumour necrosis
CC factor (TNF) alpha. The new protein is used to identify proteins which
CC bind to it (particularly to the PDZ domains) and the genes encoding them,
CC for use in the treatment of cell proliferation disorders such as cancer.
XX
SQ Sequence 2000 AA:

```

```

Query Match 100.0%; Score 5085; DB 20; Length 2000;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1005; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MLOVNSKESEFERTINIAKGNSSLGMTVSANKDGLGIVNSIHGAIISRDGRIAGDCIL 60
   |||||||
Db 996 mlgvnskesefertlnakgnsslgmtvsankdglgmivrsihngaisrdgrlaidcill 1055
QY 61 SINEESTISTVNAQARAMLRRHSLIGPDITKITYVPAHLEEFKISIGQSGRMALDIFS 120
   |||||||
Db 1056 slineestistvnaqaramlrrhsligpdikityvpaehleefkisygqsgymaldifs 1115
QY 121 SYTGNDIPELPEREGEGESELQNTAYSNMNPARRVELMREPSKLSIGISVGRMGSR 180
   |||||||
Db 1116 sytgndipelperegeegeseelqntaysnmnprrvelwrepsklsigisvgrmgstr 1175
QY 181 LSNGEVWRCGFIKRVHVEDSPAGKNGTLKPGDRIVEAPRSQSESPERAPLCSVPPPPSAR 240
   |||||||
Db 1176 lsngevwmrgfikhvledspagkngtlkpgdriveaprsqsesepkeraplcsvppppsar 1235
QY 241 AEMGSDHTOSASAKISQDVYKEDDFGYSWKNIERRGTTLGELHMIIELEKGGSLGSLA 300
   |||||||
Db 1236 aemgsdhtosasaakisdqdvkedefgyswkniergytlgelhmiielekghsglsls 1295
QY 301 GNRKDRMSVFIYIDPNGAAGKDGRLQIADLELEINGQILYGRSHQNASIITKCAPSKV 360
   |||||||
Db 1296 gnrkdrmsvfiyidpngaagkdgrlqiadelleingqillygrshqnasiltkcapskv 1355
QY 361 KIIFIRNKDAVNQMAVCPGNAVEPLPSNSENLONKETEPTVTTSDAAYDLSSFRKNVQHE 420
   |||||||
Db 1356 kifirnkdaavnqmavcpgनावेप्लप्सनेनलनकेप्टवत्सदावदलसफ्रकनव्हे 1415

```

```

OY 421 LFKDGGGLGIAISEEDTLGVTIKSLTEHGVNATDGLKLVGDQIILAVDEIVGVPIEKF 480
DB 1416 lpkdggglgialseedtlsgvllkeltehgvnadtgrllkvvgdqlavddelavvgpliekf 1475
OY 481 ISLLTAKMTVLTTHAENPDQAAPSAAGASGEKKNSQSGLAMPQSGSPPESTRNTS 540
DB 1476 ISLLTAKMTVLTTHAENPDQAAPSAAGASGEKKNSQSGLAMPQSGSPPESTRNTS 1535
OY 541 RSTPAIFASDPATCPPIIPGCEETIEISKRTGLSLIVSGSDTLTGAFIHEVEEGAA 600
DB 1536 rstpaifasdpacpplipgceetleiskrtglslivsgsdctllgafihveyeegaa 1595
OY 601 CKDGLMAGDQILEVNGIDLRKATHDEAINVLRQTPQVRVRLTYRDEAPYKEEVCDTLT 660
DB 1596 ckdgrlwagdqllenvngldlrkathdeainvlrtcpqrvtlltyrdeapykeevcdtlt 1655
OY 661 IELOKKPKGKGLSTVGRNPTGVFVSDIVKGINDPGRRLIQSGQIILVNGEDVRNNSQ 720
DB 1656 Ielgkpkpgkglstvgkrndtgvfvsdlykginadpgrlllqsgqillvngedvrnasq 1715
OY 721 EAVALLKCSLGTVLEVGRIKAGPFHSESRPSQTSQVSEGLSFTPLSGSSTSESL 780
DB 1716 eavaallkcsjgtvlevgrikagpfhseerpsqtsqvseslsftplsgsstseale 1775
OY 781 SSKKNALASERIGLRTVEMKKGPDSLGISTAGVSGPLGDVPFIAMMPTGVAAQTQ 840
DB 1836 sskknalaselrglrvemkkgpdsjglsagvsgplgdvpfiammptgyaaqtq 1835
OY 841 KLRVDRIVTTCGTSTEGTHTQAVNLKNAKSGSTEMOVAGGVSVYTGHHOEPASSL 900
DB 1836 klrvdrivlttcgstegmthtqavnlknaagslemvvaagdvsvylghhqpasasl 1895
OY 901 SFTGLTSTSIQDLDGPPOCKSITLERGPDGIFGFSIVGSGPHGDLPIYKTVPAKGA 960
DB 1896 sftgltsisifqddlgppqcksitlergpdgfgfsivgsgphgdlpiyktvfaakga 1955
OY 961 SEDGLKRGDQIIVNGOSLEGVTHEEAVATIKRKGTVTLMVLS 1005
DB 1956 sedglkrgdqllavngsllegvtheeavallkrtkgtvtlmvls 2000

```

RESULT 4
 ID AAY04733 standard; Protein: 2070 AA.
 AC AAY04733;
 DT 06-JUL-1999 (first entry)
 DE Protein containing PDZ domain from clone 38-2-1b.
 KW PDZ domain; gene expression; human umbilical vascular endothelial cell;
 KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding;
 KW cell; proliferation disorder; cancer.
 OS Homo sapiens.
 PN MO9907846-A1.
 PD 18-FEB-1999.
 PF 12-AUG-1998; 98WO-JP03603.
 PR 19-JUN-1998; 98JP-0189944.
 PR 12-AUG-1997; 97JP-0230356.
 PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
 PI Funahashi S, Miyata S;
 DR WPI: 1999-167423/14.
 DR N-PSDB: AAX29910.

PT Protein containing PDZ domain, whose expression is enhanced by TNF
 PT interactions - plays an important role in protein/protein
 PT interactions and is used for screening for proteins for use in
 PT treatment of cell proliferation disorders such as cancer

PS Claim 1; Page 167-176; 240pp; Japanese.

CC This sequence represents a new protein containing a PDZ domain encoded
 CC by the clone 38-2-1b, whose expression in human umbilical vascular
 CC endothelial cells (HUVEC) is enhanced by stimulation with tumour necrosis
 CC factor (TNF) alpha. The new protein is used to identify proteins which
 CC bind to it (particularly to the PDZ domain) and the genes encoding them,
 CC for use in the treatment of cell proliferation disorders such as cancer.

SO Sequence 2070 AA:

Query Match 99.0%; Score 5035; DB 20; Length 2070;
 Best Local Similarity 93.4%; Pred. No. 0;
 Matches 1004; Conservative 0; Mismatches 1; Indels 70; Gaps 1;

```

OY 1 MLOANSKSEFRTINIAKNSLIGMTVANNDGICMTIRSIHGAISRDRIRAGDITL 60
DB 996 mlgnskseftrtlniaknsligmtvanndgicmtirslhgaistrdriragdtl 1055
OY 61 SINESTSVTNAQARAMLRRHSLIGPDIKITYPAEHLFEFKISLGOQGRVALDIFS 120
DB 1056 slnestslvtnaqaramlrrhsligpdikitypaehlfeekslsgqgrvaldifs 1115
OY 121 SYTGNDIPELPEREGEESBELONTATSNMNOARRVRLKNEPSKLSLIVGGRMSR 180
DB 1116 sytgndipeleperegeesbelqntatynmqrrvrlwrepakslslivggrmgmr 1175
OY 181 LSNCEVMGIRIKHLESPKNGKTLKPCGRIVE----- 215
DB 1176 lsncemvgirikhlevspkngkntlpcgrive----- 1235
OY 216 -----APSQSESEPKAPLC 230
DB 1236 vfmwqslmrprkspjpsllhnlpykynfssstnpfadsldqnadapqsesepkajlc 1295
OY 231 SVPPPPSAFAEMGSDHROSSASKISODVDKEDERGYSKNNIRERYGTLNGLMIELEK 290
DB 1296 svppppsaafaemgsdhtqssaskisqdvdkederygysknnirerygtllnglmlielek 1355
OY 291 GHSGLSLAENKDRSMASVFIYGDIPNGAAGKDGRLQIADLELLEINQOILYGRSHQNAS 350
DB 1356 ghsjglslagnkdrsmasvfiygdipngaagkgdgrliqadellelengqillygrshqnas 1415
OY 351 SIICAPSKVKIIFIRNKDAVNQMAVCPGNAVEPLPSNSENLOKTEPVTTSDAAYDL 410
DB 1416 sliicapskvkiifirnkdaavnqmavcpgनावेप्लप्सनसेनलोकतेपवत्तसदायद्ल 1475
OY 411 SSFKNVQHLELPKXOGGGLIAISEEDTLGVTIKSLTEHGVNATDGLKLVGDQIILAVDE 470
DB 1476 ssfkvnqhlelpkxogggliaiseedtlsgvllkeltehgvnadtgrllkvvgdqlavdde 1535
OY 471 IVVGYPIEKFTSLTAKMTVLTTHAENPDQAAPSAAGASGEKKNSQSGLAMPQSGS 530
DB 1536 ivvgypiekfsltktamtkvltthaeenpdqaavpsaagasekknsqslampqsgs 1595
OY 531 PEPESIRNTSRSSTPAIFASDPATCPPIIPGCEETIEISKRTGLSLIVSGSDTLTGAFI 590
DB 1596 pepesirntsrsstpaifasdpacpplipgceetleiskrtglslivsgsdctllgafi 1655
OY 591 THEVYEEGAACKDGRMAGDQILEVNGIDLRKATHDEAINVLRQTPQVRVRLTYRDEAPY 650
DB 1656 theveyeegaackdgrlwagdqllenvngldlrkathdeainvlrtcpqrvtlltyrdeapy 1715
OY 651 KEEVCDFLTLELOKKPKGKGLSTVGRNPTGVFVSDIVKGINDPGRRLIQSGQIILV 710
DB 1716 keevcdtltlelgkpkpgkglstvgkrndtgvfvsdlykginadpgrlllqsgqillv 1775

```

QY 711 NGEVDNMAOEAVALLKSLGTVLEVGRIKAGPFHSRRSPQTSQVSEGLSFTFPL 770
|||||
Db 1776 ngedvnaaqaavallkslgtvtlevgrlkaqphserpsqtsqveglstftpl 1835
QY 771 SSSSTSESSLESSKKNALASEIOGLFTVMKKGPDSLSISAGVSGPLGDVPIIAM 830
|||||
Db 1836 agsstesleassakknalaseiqglrtvemkkypctdsiglsagvsgplgdvplfiam 1895
QY 831 HPTGVAOQOKIRYGRDRIYTGCTSTEGTHTHQAVALLNKASSTIMOVVAGGDSVYVG 890
1896 hptgvaagqkrlrvgrlvtlqgtslegmthcqaavllknaasglvmqvvaagdvsvvvg 1955
QY 891 HHQEPASSLSFTGLSTFISFODDLGPPOCKSTTERGPDGLGFSIVGYSPPHCDLPY 950
1956 hngepasslsftglststfisdgdlgppqckstlergpdglgfsivgyspphcdlp 2015
QY 951 VKTVEAKGASEDGRUKRGDQIIAANGSLGEGVTHEEAVAILKRTKGTVLWYLS 1005
|||||
Db 2016 vktvfakgaasedgrlkrqdgqllavngsllegvtheeavaailkrtkgtvtlwyls 2070

RESULT 5

AAV53753 standard; Protein; 2037 AA.

AAV53753;

22-FEB-2000 (first entry)

Amino acid sequence of the MMS2 protein.

Human: MMS2; MMAC1; PDZ domain; tumour suppressor; tyrosine phosphatase; scaffolding protein; cancer.

Homo sapiens.

WO958548-A1.

18-NOV-1999.

07-MAY-1999; 99WO-US09969.

08-MAY-1998; 98US-0084740.

(MYRI-) MYRIAD GENETICS INC.

Bartel PL, Tavtigian SV;

WPI: 2000-053077/04.

N-PSDB: AAZ36453.

Nucleic acids and polypeptides representing human MMS2, useful for detecting, diagnosing a predisposition to, and treating cancer -

Claim 1: Page 93-99; 112pp; English.

The present sequence represents human MMS2 protein. The MMAC1 protein binds to MMS2. The MMS2 protein has 11 post-synaptic density protein, disc-large, zc-1 (PDZ) domains and one or more of these domains interact specifically with the carboxyl terminal amino acids of MMAC1 (see AAV53754). Specifically, it appears that domain 7, 10 and 13 interact with MMAC1. Since MMS2 contains 11 PDZ domains and interacts with MMAC1, a known tumour suppressor having a region of homology with protein tyrosine phosphatases, MMS2 acts as a scaffolding protein in a common biological pathway with MMAC1. It is believed that the interaction between MMAC1 and MMS2 is required for the tumour suppressor activity of MMAC1. The MMS2 polypeptides, polynucleotides, fragments and specific or complex specific antibodies may be used for detecting cancer or a predisposition to cancer and screening for agents that may be used to treat MMS2 and/or MMAC1 related cancer. The polypeptides and polynucleotides may also be used to treat cancer.

Sequence 2037 AA;

Query Match 98.7%; Score 5018.5; DB 21; Length 2037;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 997; Conservative 5; Mismatches 3; Indels 37; Gaps 1;

QY 1 MLONVSEFEFERTINAKNSSIGMTVSANKDGLIVRSIIHGGAISDGRIGADIC 60
|||||
Db 996 mlgvnskefettlnaknsslgmtvsankdglivrsllhggaistdgrilaigdcil 1055
QY 61 SINEESTISVTNAQARAKLRHSLGDPDITITYPAEHLKEFRISLGQSGRYMALDIFS 120
1056 sineestisvtnaqaaramlrhslgdpdiktitypaeheleekfislsgqsgrymaldfis 1115
QY 121 STTGRIPELPEREEGEGSESELONTAYSNMOPRVELMREPSKLSISYGRMGSR 180
1116 sttgripelepereegegeeselnontaysnmoprvelmrepsklsisvgrmgsg 1175
QY 181 LSNGEVNRGIFIKHVEDSPACKNGTLKRGDRIVE----- 215
1176 lsngevnmrgiflkhvledspackngtlkrgdrlvevdgmdlrdashegaealrkagp 1235
QY 216 -----APSQSESEPEKAPLCVPPPPSAFAEMGSDHNTQSASKISQVDXED 263
1236 vfmvgslnrrprapgseseepetkplcsyppppsa faemgdsghgsasaklsqdvdked 1295
QY 264 EFGYSKKNIRERYGTLTGELHMIETLEKSHGLSLAGLNKDRSRMSVFTVGDIPNGAOK 323
1296 efgyskknirerygtltgelhmieltekshgslaglnkdrsrmsvftvgdipngaok 1355
QY 324 DGRLOIADELLEINCOILYGRSHONASSIICAPSVKIIIFIRNKAVNMVCPENAYE 383
1356 dgrlgiadelleinglilygrshonassllkapsvkvllifirnkavnmvcpnaye 1415
QY 384 PLPSNSENLOKNETEPTVTTSAAVDLSSEFKNOHLEPKRDOGLGIAISEEDTLGVI 443
1416 plpsnseelnokneteptvttsdaavdlssfknohlepkrdoglgiaiseedtlsgvll 1475
QY 444 KSLTEHGAATDGRKLVGQOIIAVIDEIVGYPTEKFISLTKAKTFVLTTHAENPDSO 503
1476 ksltehgaatdgrlkvqgqllavddelivgyplekfisltkaktvltlhaenpdsq 1535
QY 504 AVPSAAGAASGEKKNSSOSLWMPQSGPEPESIRNRSSTPAIFASDPATCPIIGCEP 563
1536 avpsaagaasgekknsoslmwpqsgpepestirnrststpaifasdpatcpiigpcc 1595
QY 564 TIEISKRTGLSLIVGSDTLGAFIIEHYEEGAACKDGRLMAGDQIILEVNGIDLRA 623
1596 tieiskrtglslivgsdtilgafiihyeeegaackdgrlmagdqiilevngidlrlra 1655
QY 624 THDEAINVRQTPQRVRLTLRYDEAPYKEEVCDDTLTELQKPKGCLSLYGRKNDPG 683
1656 thdeainvrqtpqrvrltlrydeapykeevcdtltelqkpkkgclslvgrkndpg 1715
QY 684 VFVSDIVKGCIGADPDGRLLQGDQIILVNGEDVYRNMAOEAVALLKSLGTVLEVGRIKA 743
1716 vfvsvdivkcgigadpdrllqgdqiilvngedvyrnaaqaavallkslgtvtlevgrika 1775
QY 744 GFPHSERRSPQTSQVSEGLSFTFPLSGSSSTSESSLESSKKNALASEIOGLRTVEMKKG 803
1776 gfphseerrpsqtsqveglstftplsgssstselesskknalaseiqglrtvemkkg 1835
QY 804 PDSLSISAGVSGPLGDVPIFIAMHRTGVAAOQOKIRYGRDRIYTGCTSTEGTHTHQ 863
1836 pdslsisagvsgplgdvplfiammhrtgvaagqkrlrvgrlvtlqgtslegmthc 1895
QY 864 AVNLKKNAGSITEMQVAVAGDVSVYVGHQEPASSLSFTGLSTFISFODDLGPPOCKSI 923
1896 avnlkknasgslmqvavagdvsvvghqepasslsftglststfisdgdlgppqcksi 1955
QY 924 TLERGPDGLGFSIVGYSPPHCDLPYVKTVEAKGASEDGRUKRGDQIIAANGSLGEGV 983
1956 tlergpdglgfsivgyspphcdlpvktvfkgaasedgrlkrqdgqllavngsllegv 2015

OY 984 THEEAVALIKRTKGTVTLMVLS 1005
 |||||
 DB 2016 theeavallkrtkgtvtlmvls 2037

RESULT 6
 AAB01383
 ID AAB01383 standard; Protein: 856 AA.
 XX
 AC AAB01383;
 XX
 DT 20-OCT-2000 (first entry)
 XX
 DE Neuron-associated protein.

XX Neuron associated protein: NEUP; neurological disorder; epilepsy;
 KW ischemic cerebrovascular disease; stroke; cerebral neoplasm;
 KW Alzheimer's disease; Pick's disease; Huntington's disease;
 KW dementia; Parkinson's disease; demyelinating disease; meningitis;
 KW prion disease; Kuru; Creutzfeldt-Jakob disease; neurofibromatosis;
 KW cerebral palsy; muscular dystrophy; central nervous system; CNS;
 KW peripheral nervous system; PNS; myopathy; schizophrenia;
 KW actinic keratosis; arteriosclerosis; atherosclerosis; burstitis;
 KW cirrhosis; hepatitis; mixed connective tissue disease; MCTD;
 KW myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;
 KW autoimmune disease; inflammation; acquired immunodeficiency syndrome;
 KW AIDS; Addison's disease; adult respiratory distress syndrome;
 KW allergy; ankylosing spondylitis; amyloidosis; anaemia; asthma;
 KW Werner syndrome; trauma; human.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Domain 1..65
 FT Modified-site 18 /label- PDZ domain signature
 FT Modified-site 20 /note- "Potential glycosylation site"
 FT Modified-site 69 /note- "Potential phosphorylation site"
 FT Modified-site 71 /note- "Potential phosphorylation site"
 FT Modified-site 73 /note- "Potential phosphorylation site"
 FT Modified-site 118 /note- "Potential phosphorylation site"
 FT Modified-site 119 /note- "Potential phosphorylation site"
 FT Modified-site 136..218 /note- "Potential phosphorylation site"
 FT Domain 144..147
 FT Modified-site 144..147 /label- PDZ domain signature
 FT Modified-site 199 /note- "Glycosaminoglycan attachment site"
 FT Modified-site 238 /note- "Potential glycosylation site"
 FT Modified-site 253 /note- "Potential phosphorylation site"
 FT Modified-site 263 /note- "Potential phosphorylation site"
 FT Modified-site 269..349 /note- "Potential phosphorylation site"
 FT Domain 284 /label- PDZ domain signature
 FT Modified-site 284 /note- "Potential phosphorylation site"
 FT Modified-site 296 /note- "Potential phosphorylation site"
 FT Modified-site 337 /note- "Potential phosphorylation site"
 FT Modified-site 341 /note- "Potential phosphorylation site"
 FT Modified-site 369 /note- "Potential phosphorylation site"

FT Modified-site 370 /note- "Potential glycosylation site"
 FT Modified-site 386 /note- "Potential phosphorylation site"
 FT Modified-site 389 /note- "Potential phosphorylation site"
 FT Modified-site 390 /note- "Potential glycosylation site"
 FT Modified-site 414 /note- "Potential phosphorylation site"
 FT Modified-site 415..497 /note- "Potential phosphorylation site"
 FT Domain 475 /label- PDZ domain signature
 FT Modified-site 475 /note- "Potential phosphorylation site"
 FT Domain 511..592 /label- PDZ domain signature
 FT Modified-site 531 /note- "Potential glycosylation site"
 FT Modified-site 568 /note- "Potential glycosylation site"
 FT Modified-site 599 /note- "Potential glycosylation site"
 FT Modified-site 604 /note- "Potential phosphorylation site"
 FT Modified-site 625 /note- "Potential phosphorylation site"
 FT Modified-site 633 /note- "Potential phosphorylation site"
 FT Modified-site 634 /note- "Potential phosphorylation site"
 FT Modified-site 648..733 /note- "Potential phosphorylation site"
 FT Domain 690 /label- PDZ domain signature
 FT Modified-site 705 /note- "Potential phosphorylation site"
 FT Modified-site 721 /note- "Potential phosphorylation site"
 FT Domain 773..856 /note- "Potential glycosylation site"
 FT Modified-site 819..821 /label- PDZ domain signature
 FT Region 835 /label- Cell attachment sequence
 FT Modified-site 835 /note- "Potential phosphorylation site"

MO200034477-A2.

15-JUN-2000.

10-DEC-1999; 99MO-US30408.

11-DEC-1998; 98US-0210083.

11-DEC-1998; 98US-9123456.

09-FEB-1999; 99US-0119365.

16-MAR-1999; 99US-0124687.

(IMCY-) INCYTE PHARM INC.

Tang YT, Yue H, Baughn MR, Hillman JL, Lai P, Au-Young J, Yang J;

Lu DAM, Azimzal Y;

WPI: 2000-423423/36.

N-PSDB; AAA47424.

New human neuron-associated proteins and polynucleotides encoding them,

useful for diagnosis, treatment and prevention of cell proliferative disorders including cancer, neuronal and neurological disorders

Claim 1; Page 105-108; 145pp; English.

Human neuron-associated proteins (NEUP) can be used for for treating or preventing a disorder associated with decreased


```

Db 1322 edavngmavtp-----fypssasp-----s1-----edqsg 1348
Oy 428 LGAISEEDTLISGVIIKSLTEHGVAATDGRKLVGQDIIAIVDEIIVGPIEKFISLTKTA 487
Db 1349 leplasee-----dgslevglkqlpesefk1a-----vsgmkqg 1383
Oy 488 KMTVLTTHAENPDSONAVPSAAGASGEKKNSOSLAMPQSSPEPESTRNRSSTPAI 547
Db 1384 kypktvstf-----ssqelplapass-----yhsldadftgyggfagplav----- 1423
Oy 548 FASDPATPIIFGCEETTEISKRTGLSIYGSGLTLLGATTHIEVEEGAACDGRML 607
Db 1424 ---dpactpivpgemileistkgrsglglstivgkdpdnalvheveegaaadgrlw 1480
Oy 608 AGDQILEVNGIDLKRAKTHDEAINVLRQTPQRVRLTYRDEAPYKEEVCDTITLQKKP 667
Db 1481 agdqllenvngvdlrnsheeaaltalrqpqrlyvrydeahyrddeanlelfpvdlqkka 1540
Oy 668 GKGGLSTVIGKRNDFVFSDIYKGIADPDGRLLGDDIILLNGEDVNRNASEAVNAL 727
Db 1541 gfglslstivgkngsvlstdivkgsaadlgrllygdqllsvngedmrnasqetvatll 1600
Oy 728 KCSLGTVTLEVGRIKAGPFRHSERPSQTSQVSEGLSFTFP-----LSGSSTSESL 779
Db 1601 kaagllvqlelgrlragvstsaaritsqngsqgshaschpsafavltglqlnlygktrv 1660
Oy 780 ESSSKNMLASIEIGLRYEYKMKGPDLISIGISAGVSPCLDVLFIAMHPTGYAOT 839
Db 1661 sdpskns-gldmep-rtyelrrelsdalglslagrgsplygdvfilamqasgvaar 1718
Oy 840 OKLRAGDRIVTICGTSTEGMTHQAVNLKKNAGSITEMOVNKGVSIVY-----TGH 891
Db 1719 qklkvgdrlvsingpldglshadvnllknaaygrlllgvavdtnlssaaqlennstgy 1778
Oy 892 HDEPASSLSFTGLTSTSIPODLGPPCKSITLERGPDGLGFSIVGCGSPHGDLPITY 951
Db 1779 h-----lgsptlaehpedteppkiltleksgelgfsivgygspghdplpyv 1828
Oy 952 KRVFAKGAASEGRKLRQDQIIAVNGQSLEGTHEEAVAILKRTGTYLAML 1004
Db 1829 klvfakgaadgrlkrqgdlavngeltgylhegavallkhgrtyvlv 1881

RESULT 9
AAV74150
ID AAV74150 standard; Protein: 318 AA.
AC AAV74150;
DT 14-MAR-2000 (first entry)
DE Human prostate tumor EST fragment derived protein #337.
KW Pancreas; tumor; EST; expressed sequence tag; human; cytostatic; treatment.
OS Homo sapiens.
PN DE19820190-A1.
PD 04-NOV-1999.
PR 28-APR-1998; 98DE-1020190.
PR 28-APR-1998; 98DE-1020190.
PR 28-APR-1998; 98DE-1020190.
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX WPI; 1999-621386/54.
XX N-PSDB; AA52969.

```

```

XX New human nucleic acid sequences from pancreatic tumors, and related
XX proteins
XX Claim 23; Page 449-450; 502pp; German.
PS This invention describes novel polypeptides and their encoding nucleic
XX acids derived from human pancreatic tumor tissue which have cytostatic
XX activity. The sequences are also useful in producing pharmaceutical
CC compositions for treatment of pancreatic tumors. AAV73814-Y74252
CC represent protein fragments encoded by the human pancreatic tumor CDNA
CC library derived expressed sequence tag (EST) sequences represented in
XX AA52858-453014.
XX
XX Sequence 318 AA:
XX
Query Match 29.5%; Score 1501; DB 20; Length 318;
Best Local Similarity 96.7%; Pred. No. 7,8e-118;
Matches 297; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
Oy 422 PKDGGCIGAISEEDTLISGVIIKSLTEHGVAATDGRKLVGQDIIAIVDEIIVGPIEKF 481
Db 9 prrggllglalseedtlsgvllksltehgvaatdgrllkvgdqlavddelivvgylekfl 68
Oy 482 SILTKAKMTVKTTHAENPDSONAVPSAAGASGEKKNSOSLAMPQSSPEPESTRNRSR 541
Db 69 sltkakmtvktlhaenpdsoavpsaagaasgekknsqslmpqsgspesestrnrsr 128
Oy 542 SSTPAIFASDPATPIIFGCEETTEISKRTGLSIYGSGLTLLGATTHIEVEEGAAC 601
Db 129 sstplfssdpactpilpgcetteliskrgtglslivgsdltllgalllhevyegaac 188
Oy 602 KDGRLMAGDQILEVNGIDLKRAKTHDEAINVLRQTPQRVRLTYRDEAPYKEEVCDTITI 661
Db 189 kdgrlmagdqllenvngidlkrkthdeainvlrqpqrlyltyrdeapykeevcdtltl 248
Oy 662 ELQKKRPGGLSIYKRNDFVFSDIYKGIADPDGRLLGDDIILLNGEDVNRNASE 721
Db 249 elqkkrpgglslivgkngsvlstdivkgsaadgrllmgdqlimvngedvnrnaqge 308
Oy 722 AVALALK 728
Db 309 avavalk 315

RESULT 10
AAV04738
ID AAV04738 standard; Protein: 251 AA.
AC AAV04738;
DT 06-JUL-1999 (first entry)
DE PDZ domain-containing protein gene encoded by clone FH750.
KW PDZ domain; gene expression; human umbilical vascular endothelial cell;
KW HUVEC; stimulation; tumor necrosis factor; TNF; protein binding; PCR;
KW cell; proliferation disorder; cancer; primer; amplification.
OS Homo sapiens.
PN WO9907846-A1.
PD 18-FEB-1999.
PR 12-AUG-1998; 98WO-JP03603.
PR 19-JUN-1998; 98JP-0189944.
PR 12-AUG-1997; 97JP-0230356.
XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX

```

Query Match	22.2%;	Score 1127;	DB 20;	Length 251;
Best Local Similarity	100.0%;	Pred. No. 1.6e-86;		
Matches 223; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

RESULT	11
AAV04739	
ID	AAV04739 standard; Protein; 272 AA

DR	N-PSDB; AAX29970.
XX	
PT	Protein containing PDZ domain, whose expression is enhanced by TNF stimulation - plays an important role in protein/protein interactions and is used for screening for proteins for use in treatment of cell proliferation disorders such as cancer
PT	
XX	
PS	Example 7; Page 152-154; 240pp; Japanese.
XX	
CC	This sequence represents a new protein containing PDZ domains encoded by clone FH850, whose expression in human umbilical vascular endothelial cells (FHVEC) is enhanced by stimulation with tumour necrosis factor (TNF) alpha. The new protein is used to identify proteins which bind to it (particularly to the PDZ domains) and the genes encoding them, for use in the treatment of cell proliferation disorders such as cancer.
CC	
CC	
CC	
CC	
SQ	Sequence 272 AA;
XX	
Query Match	21.4%; Score 1089; DB 20; Length 272;
Best Local Similarity	100.0%; Pred. No. 2.8e-83;
Matches 215; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 MIQNVSKSFERTTIAAGNSLGLMTVSNKDGCLMYRSTTHGCAISRDGAIAGDCTL 60
Dd	29 mlgvsksfertlinlakgnslgmvtaankglmivrsilhgsalardgilaigdcil 88
OY	61 SINESTSVTAQAARALRRHSILGPDIKITYPVAEHLPEFKISLGQSGRWMLDIFS 120
Dd	89 sinestlsyvnagarnlrhrsilgpdtkitlypaehneekislqggsgryvmldifs 148
OY	121 SYTGNDIPELPEREGEESELEQTAVYSNMNQPRVELMRPESKSLGISIVGNMGSR 180
Dd	149 sytgndipelpereeggeeselnqltaysnmnqprrvelmrrepskslglislvgrmgssr 208
OY	181 LSNCEVMRGIFIKHYLEDSPAKGNTLKPGDRIVE 215
Dd	209 lsngevrmrgifikhvledspakngntlkpgdrive 243
RESULT 12	
AAY04740	
ID	AAY04740 standard; Protein; 319 AA.
XX	
AC	AAY04740;
XX	
DT	06-JUL-1999 (first entry)
XX	
DE	PDZ domain-containing protein gene encoded by clone FH950.
XX	
KW	PDZ domain; gene expression; human umbilical vascular endothelial cell; HUVEC; stimulation; tumour necrosis factor; TNF; protein binding; PCR; cell; proliferation disorder; cancer; primer; amplification.
XX	
OS	Homo sapiens.
XX	
PN	WO9907846-A1.
XX	
PD	18-FEB-1999.
XX	
PF	12-AUG-1998; 98WO-JP03603.
XX	
PR	19-JUN-1998; 98JP-0189944.
XX	
PR	12-AUG-1997; 97JP-0230356.
XX	
PA	(CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX	
PI	Funahashi S, Miyata S;
XX	
WI	WPI; 1999-167423/14.
XX	
DR	N-PSDB; AAX29971.
XX	
PT	Protein containing PDZ domain, whose expression is enhanced by TNF

stimulation - plays an important role in protein/protein interactions and is used for screening for proteins for use in treatment of cell proliferation disorders such as cancer

Example 7, Page 155-158; 240pp; Japanese.

This sequence represents a new protein containing PDZ domains encoded by clone FH950, whose expression in human umbilical vascular endothelial cells (HUVEC) is enhanced by stimulation with tumour necrosis factor (TNF) alpha. The new protein is used to identify proteins which bind to it (particularly to the PDZ domains) and the genes encoding them, for use in the treatment of cell proliferation disorders such as cancer.

Sequence 319 AA;

Query Match 21.4%; Score 1089; DB 20; Length 319;
Best Local Similarity 100.0%; Pred. No. 3,7e-83;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLYNYSKSEFERTINANGNSSLGWTVSANKDGLMIVRSIIHGCAISPDGRIAGDCTL 60
DB 29 mlynyksesferltlnakynsslgmtvsankdglmivrsllhggalsrdgrlaldcll 88
QY 61 SINESTISVTNAQARAMLRRHSLIGPDIKITYVPAEHLKEFKISLGQSGRWALDIFS 120
DB 89 sineestlsvtnaqaramlrrhsligpdikityvpaehleekfkslgsgsgvymaldifs 148
QY 121 SYTGNDIPELPEREGEGESELQNTAVSNMNPRLVRLREPSKSLGISIVGSGMGSR 180
DB 149 sytgndipeelperegeegeseelqntaysnmnprrvlwrepskslgislvgrgmgr 208
QY 181 LSNGEVMRGIFIKHVLDESPACKNGTLKPGDRIVE 215
DB 209 lsngevrmrgiflkhvledspagknltlpgdrive 243

RESULT 13
AAV04734

ID AAV04734 standard; Protein; 1239 AA.

AAV04734;

06-JUL-1999 (first entry)

Protein containing PDZ domain from clone 38-2-1c.

PDZ domain; gene expression; human umbilical vascular endothelial cell;
HUVEC; stimulation; tumour necrosis factor; TNF; protein binding;
cell; proliferation disorder; cancer.

Homo sapiens.

WO9907846-A1.

18-FEB-1999.

12-AUG-1998; 98WO-0P03603.

19-JUN-1998; 98JP-0189944.

12-AUG-1997; 97JP-0230356.

(CHUGAI) CHUGAI RES INST MOLECULAR MEDICINE INC.

Funahashi S, Miyata S;

WPI: 1999-167423/14.

N-PSDB; AAX29911.

Protein containing PDZ domain, whose expression is enhanced by TNF stimulation - plays an important role in protein/protein interactions and is used for screening for proteins for use in treatment of cell proliferation disorders such as cancer

Claim 1, Page 177-182; 240pp; Japanese.

This sequence represents a new protein containing a PDZ domain encoded by the clone 38-2-1c, whose expression in human umbilical vascular endothelial cells (HUVEC) is enhanced by stimulation with tumour necrosis factor (TNF) alpha. The new protein is used to identify proteins which bind to it (particularly to the PDZ domains) and the genes encoding them, for use in the treatment of cell proliferation disorders such as cancer.

Sequence 1239 AA;

Query Match 21.4%; Score 1089; DB 20; Length 1239;
Best Local Similarity 100.0%; Pred. No. 3,5e-82;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLYNYSKSEFERTINANGNSSLGWTVSANKDGLMIVRSIIHGCAISPDGRIAGDCTL 60
DB 996 mlynyksesferltlnakynsslgmtvsankdglmivrsllhggalsrdgrlaldcll 1055
QY 61 SINESTISVTNAQARAMLRRHSLIGPDIKITYVPAEHLKEFKISLGQSGRWALDIFS 120
DB 1056 sineestlsvtnaqaramlrrhsligpdikityvpaehleekfkslgsgsgvymaldifs 1115
QY 121 SYTGNDIPELPEREGEGESELQNTAVSNMNPRLVRLREPSKSLGISIVGSGMGSR 180
DB 1116 sytgndipeelperegeegeseelqntaysnmnprrvlwrepskslgislvgrgmgr 1175
QY 181 LSNGEVMRGIFIKHVLDESPACKNGTLKPGDRIVE 215
DB 1176 lsngevrmrgiflkhvledspagknltlpgdrive 1210

RESULT 14
AAV74151

ID AAV74151 standard; Protein; 206 AA.

AAV74151;

14-MAR-2000 (first entry)

Human prostate tumor EST fragment derived protein #338.

Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;
treatment.

Homo sapiens.

DE19820190-A1.

04-NOV-1999.

28-APR-1998; 98DE-1020190.

28-APR-1998; 98DE-1020190.

(META-) METAGEN GES GENOMFORSCHUNG MBH.

Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;

WPI: 1999-621386/54.

N-PSDB; AAZ52969.

New human nucleic acid sequences from pancreatic tumors, and related proteins

Claim 23; Page 450; 502pp; German.

This invention describes novel polypeptides and their encoding nucleic acids derived from human pancreatic tumor tissue which have cytostatic activity. The sequences are also useful in producing pharmaceutical compositions for treatment of pancreatic tumors. AAV73814-VY4252

Query Match	11.5%;	Score 583;	DB 21;	length 632;
Best Local Similarity	25.6%;	Pred. No. 4.4e-40;		
Matches 195;	Conservative 119;	Mismatches 254;	Indels 194;	Gaps 21.

Search completed: July 12, 2001, 14:40:37
Job time: 200 sec

Fri Jul 13 15:00:10 2001

us-09-502-698-2.rag

Page 14

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw-model

Run on: July 12, 2001, 14:41:51 ; Search time 40.62 Seconds
(without alignments)
1884.673 Million cell updates/sec

Title: US-09-502-698-2

Perfect score: 5085
Sequence: 1 MGNVSKESFERTINIAKGN.....EENVALTKRKGVTLVNLVS 1005

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR.68:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4413.5	86.8	2054	2	T46612 multi PDZ domain p
2	4314.5	84.8	2055	2	T30259 multiple PDZ domai
3	814	16.0	2172	2	T20145 hypothetical prote
4	589	11.6	728	2	T09457 numb-binding prote
5	586	11.5	628	2	T09458 numb-binding prote
6	534	10.5	1012	2	T23160 hypothetical prote
7	523	10.3	2450	2	S71625 protein-tyrosine-p
8	509.5	10.0	2466	2	T67629 protein-tyrosine p
9	501.5	9.9	2294	2	T67630 protein-tyrosine p
10	494.5	9.7	2490	1	A54971 protein-tyrosine-p
11	453.5	8.9	1256	2	JR0209 brain-specific ang
12	407.5	8.0	852	2	T10811 channel associated
13	399.5	7.9	1131	2	T15617 hypothetical prote
14	397.5	7.8	870	2	G01974 channel associated
15	395	7.8	1277	2	T14152 synaptic scaffoldi
16	385	7.6	1171	2	T42372 probable guanylate
17	384	7.6	767	2	T09599 postsynaptic densi
18	383	7.5	724	2	JH0800 postsynaptic densi
19	380	7.5	720	2	A45436 synapse-associated
20	379	7.5	960	1	A39651 discs-large tumor
21	378.5	7.4	911	2	T16552 synapse-associated
22	377	7.4	904	2	T18757 homolog of Drosoph
23	377	7.4	926	2	T32756 AMPA glutamate rec
24	357	7.0	1112	2	T32733 atypical protein k
25	352	6.9	1337	2	T13948 bacloxa gene prote
26	348	6.8	1464	2	T13716 cell polarity prot
27	264	5.2	1360	2	T34302 tight junction pro
28	246	4.8	1163	2	JF0366 tight junction pro
29	243	4.8	1736	2	A47747

30	237.5	4.7	1745	2	A46431 tight junction-ass
31	228	4.5	1116	2	I54378 gene X104 protein
32	216	4.2	1367	2	S60315 PSD-95-related pro
33	210	4.1	1367	2	T13703 tama protein - fru
34	197	3.9	505	2	S62894 alpha-syntrophin -
35	194.5	3.8	126	2	I81210 tyrosine phosphata
36	190.5	3.7	1829	2	T41751 l-afadin - rat
37	188	3.7	1281	2	T00346 hypothetical prote
38	187	3.7	723	2	T14765 hypothetical prote
39	187	3.7	1095	2	T43275 neurabin - rat
40	187	3.7	1663	2	T42092 s-afadin - rat
41	187	3.7	1893	2	A56158 eye development pr
42	186.5	3.7	1117	2	I81209 tyrosine phosphata
43	186.5	3.7	1666	2	T43169 hypothetical prote
44	186.5	3.7	5627	2	C83359 hypothetical prote
45	185.5	3.6	488	2	I51379 syntrophin - Pacif

ALIGNMENTS

RESULT 1

T46612 multi PDZ domain protein 1 - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 21-Jul-2000

C:Accession: T46612

R:Fullmer, C.; Schmeck, K.; Flyge, A.; Lubbert, H.

A:Title: Cloning and characterization of MUPP1, a novel PDZ domain protein.

A:Reference number: 223104; MUID:98196865

A:Accession: T46612

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-2054 <UDL>

A:Cross-references: EMBL:AJ001320; NID:g2959978; PIDN:CAA04681.1; PID:g2959979

A:Experimental source: brain

C:Genetics:

A:Gene: MUPP1

Query Match	86.8%	Score 4413.5;	DB 2;	Length 2054;
Best local similarity	81.5%	Pred. No. 5.1e-244;		
Matches	875;	Conservative	64;	Mismatches 62; Indels 73; Gaps 2;
QY	2	LGNVSKESFERTINIAKGN	SLGNTVSAKNDGLGNVSIHGAISRDGRIAGDCILS	61
DB	984	LQSMQEAFFERTVTAKSSSLGNTVSAKNDGLGIVRSIIHGAIISRDGRIAGDCILS		1043
QY	62	INEESTISVTNAQARAMLRHSLSIGPDKITVPAHEHLEEFKISLGQOSGRVMAIDFSS		1103
DB	1044	INEESTISVTNAQARAMLRHSLSIGPDKITVPAHEHLEEFVSGQAGGMALDIFSS		1103
QY	122	YTGRIPELPEREGEGESELQNTAVSNMNPRAVELMREPSKSLGISIVGRCMGSR		181
DB	1104	YTGRIPELPEREGEGESELQNTAVSNMNPRAVELMREPSKSLGISIVGRCMGSR		1163
QY	182	SNGEVVRGIFIKVHVEDSPAGKNGTLKPEDRIVE		215
DB	1164	SNGEVVRGIFIKVHVEDSPAGKNGTLKPEDRIVE		1223
QY	216	APSOSESEPEKAPICS		231
DB	1224	FWQSTIVNRPKSPPLSLPHSLYPKCSSTNPPAESLQITSDKAPSOSESEPEKAPICS		1283
QY	232	VPPPPSAFAEMGSDHTOSASAKISQDYDKEDDFGYSKKNIRERYGTLTGLHMLEKRG		291
DB	1284	VPSSPSVSEKSSDYADPATTVADEKDEDFGYSKKNIRERYGTLTGLHMLEKRG		1343
QY	232	HSGGLSLAGNKDRSRMSVFIYIDPNGAAGKGLQIADLELENGOITLGRSQNASS		351
DB	1344	HSGGLSLAGNKDRSRMSVFIYIDPNGAAGKGLQIADLELENGOITLGRSQNASS		1403

C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T20145; T25009
 R:Stulston, J.
 submitted to the EMBL Data Library, November 1994
 A:Reference number: Z19229
 A:Accession: T20145
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2172 <MTL>
 A:Cross-references: EMBL:Z46792; PIDN:CAA6769.1; GSPDB:GN00020; CESP:C52A11.4
 A:Experimental source: clone C52A11
 R:Colles, L.
 submitted to the EMBL Data Library, November 1994
 A:Accession: T25009
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2172 <MT2>
 A:Cross-references: EMBL:Z46795; PIDN:CAA6769.1; GSPDB:GN00020; CESP:C52A11.4
 A:Experimental source: clone T19E10
 C:Genetics:
 A:Gene: CESP:C52A11.4
 A:Map position: 2
 A:Introns: 27/2; 78/1; 111/3; 156/3; 208/1; 230/3; 290/2; 341/1; 389/3; 431/1; 461/3; 63
 24/3; 1747/1; 1946/3; 2049/3; 2096/3; 2145/3

Query Match 16.0%; Score 814; DB 2; Length 2172;
 Best Local Similarity 20.9%; Pred. No. 4e-38;
 Matches 307; Conservative 137; Mismatches 259; Indels 768; Gaps 33;

DB 11 ETTIANGNSLGMTVSNKND-GI-GMTVRSIIHGASRGRRAIGCISTINEESTI 68
 ETTIANGNSLGMTVSNKND-GI-GMTVRSIIHGASRGRRAIGCISTINEESTI 68
 DB 859 ETTIANGNSLGMTVSNKND-GI-GMTVRSIIHGASRGRRAIGCISTINEESTI 68
 ETTIANGNSLGMTVSNKND-GI-GMTVRSIIHGASRGRRAIGCISTINEESTI 68
 DB 69 SVTNAQARMLRRHLIGPDKITVVP----- 95
 SVTNAQARMLRRHLIGPDKITVVP----- 95
 DB 919 NNTNQAARILKRTNLVGFCAVNTYITSDATKWKERPORPSESSPIINRLSPKVFKE 978
 NNTNQAARILKRTNLVGFCAVNTYITSDATKWKERPORPSESSPIINRLSPKVFKE 978
 DB 96 -----AEHLERKISLG-----QSGRYM 114
 -----AEHLERKISLG-----QSGRYM 114
 DB 979 YRSPMROESQSKTEMTDETEAPSIMTDSMEHIKTFDLAEGSSRSHHDEQEVNRMS 1038
 YRSPMROESQSKTEMTDETEAPSIMTDSMEHIKTFDLAEGSSRSHHDEQEVNRMS 1038
 DB 115 AL-----DIF 119
 AL-----DIF 119
 DB 1039 RLIDGVEVDENVNLIKKAITDATTIELRLVLRKTKDMSNCRERLESPLPPPEVL 1098
 RLIDGVEVDENVNLIKKAITDATTIELRLVLRKTKDMSNCRERLESPLPPPEVL 1098
 DB 120 SS-----YTGRI-----PELPE 132
 SS-----YTGRI-----PELPE 132
 DB 1099 SSPKSPVAVQPTPRDLEEVLTSTASSLEHSGORTSOLHILSTEEVLOATPPSPFE 1158
 SSPKSPVAVQPTPRDLEEVLTSTASSLEHSGORTSOLHILSTEEVLOATPPSPFE 1158
 DB 133 REEGE----- 137
 REEGE----- 137
 DB 1159 NKSSEVPSPISPGIKAGEVTAPEIEIYVKEQAEVNDRAETAATGAEEVATSTPAEAI 1218
 NKSSEVPSPISPGIKAGEVTAPEIEIYVKEQAEVNDRAETAATGAEEVATSTPAEAI 1218
 DB 138 -----GESELOTA-----Y 148
 -----GESELOTA-----Y 148
 DB 1219 GNSKDESTTISISOQSVGLQTALNSTEVNSNMSRVTSTPTSGES-LQNAQOLVRS 1277
 GNSKDESTTISISOQSVGLQTALNSTEVNSNMSRVTSTPTSGES-LQNAQOLVRS 1277
 DB 149 SNMNPRLRVETLRSPKSLGISIVGR-----GMSRLNGEVMRGIFIKHVLVEDSPAGKN 204
 SNMNPRLRVETLRSPKSLGISIVGR-----GMSRLNGEVMRGIFIKHVLVEDSPAGKN 204
 DB 1278 KYWGAFARTVTLVREPNKSGISIVGGRVYSOKGGLPGTGNVCGIFIKSVLPNSPAGNS 1337
 KYWGAFARTVTLVREPNKSGISIVGGRVYSOKGGLPGTGNVCGIFIKSVLPNSPAGNS 1337
 DB 205 GTLKDGDRIVEAP----- 217
 GTLKDGDRIVEAP----- 217
 DB 1338 GQNMNDRIVSYNDVLDRADEQAVNAIKNASNPVRFVLSLHTNQNMINSASNTVG 1397
 GQNMNDRIVSYNDVLDRADEQAVNAIKNASNPVRFVLSLHTNQNMINSASNTVG 1397
 DB 218 -----SSESEEPKAPLCSPP-----PPPS-----AFALMGSHQDS 250
 -----SSESEEPKAPLCSPP-----PPPS-----AFALMGSHQDS 250
 DB 1398 SVRFENAKFEELPPALVTPPLKPMISSSGSSTSKPAMNPPPSISTTTTTSMESESKBE 1457
 SVRFENAKFEELPPALVTPPLKPMISSSGSSTSKPAMNPPPSISTTTTTSMESESKBE 1457
 DB 251 SASKISQDVDEKED----- 263
 SASKISQDVDEKED----- 263

DB 1458 EATSSSPDIOGRNTYKRRSMOYERQEPANPVETSIYKERTPPRIKSAKSSQKESND 1517
 EATSSSPDIOGRNTYKRRSMOYERQEPANPVETSIYKERTPPRIKSAKSSQKESND 1517
 DB 264 ----- 263
 ----- 263
 DB 1518 RKEIKKOKSVREMSVESKKSIVRSIKKHROESTIIKSPNETABLDIVDSSETHDE 1577
 RKEIKKOKSVREMSVESKKSIVRSIKKHROESTIIKSPNETABLDIVDSSETHDE 1577
 DB 264 -----EGYSWKNIIRKRYGLTDEL 283
 -----EGYSWKNIIRKRYGLTDEL 283
 DB 1578 PQAMSPSTSPDTRDAMRALGIDDDSAFQIKNDGEPSKFFYAGRIERRYDSDGSEL 1637
 PQAMSPSTSPDTRDAMRALGIDDDSAFQIKNDGEPSKFFYAGRIERRYDSDGSEL 1637
 DB 284 HMELEKSHGLSLAGNKRMSRVFIVGIDPGAAGKDRLOIADLEINQILYG 343
 HMELEKSHGLSLAGNKRMSRVFIVGIDPGAAGKDRLOIADLEINQILYG 343
 DB 1638 VLVACERPDGGLISLAKNKRDKONVFNVRPSCPILA-----IRPGELLEINRLINK 1693
 VLVACERPDGGLISLAKNKRDKONVFNVRPSCPILA-----IRPGELLEINRLINK 1693
 DB 344 RSHONASSITK-CAP--SKVKIIFIRNKDAVQNAVCQNAVEPLPSENLOKRETEPT 400
 RSHONASSITK-CAP--SKVKIIFIRNKDAVQNAVCQNAVEPLPSENLOKRETEPT 400
 DB 1694 ISHVAASAVRECCQOHONIEIVLRNRGALKSTA-----PQMTTEGLEN----- 1740
 ISHVAASAVRECCQOHONIEIVLRNRGALKSTA-----PQMTTEGLEN----- 1740
 DB 401 VTTSDAAVDLSFKNVQHLPLKDGGLGIALSEEDTISGVIRKSLTEHGVATDGRUKV 460
 VTTSDAAVDLSFKNVQHLPLKDGGLGIALSEEDTISGVIRKSLTEHGVATDGRUKV 460
 DB 1741 -----AAPNOEELS----- 1749
 -----AAPNOEELS----- 1749
 DB 461 GQIILAVDEIVGVPIKEFISLTKAKMTVLTITHAENPDSQAVPSAAGAAGEKKNSS 520
 GQIILAVDEIVGVPIKEFISLTKAKMTVLTITHAENPDSQAVPSAAGAAGEKKNSS 520
 DB 1750 -----RKKSFS 1755
 -----RKKSFS 1755
 DB 521 QSLMVPQSGSPESIRNTSRSTPAIFASDPATCPIIPGCTTIEISKGTGLSTVG 580
 QSLMVPQSGSPESIRNTSRSTPAIFASDPATCPIIPGCTTIEISKGTGLSTVG 580
 DB 1756 QR-----RTQAIEN-----GRETMIEIDKDGKGLSTVG 1785
 QR-----RTQAIEN-----GRETMIEIDKDGKGLSTVG 1785
 DB 581 GSDTLGAFIIEHYEEGACCKDRLMAGDQILEVNGIDLKRAETHDEAINVROTPQVR 640
 GSDTLGAFIIEHYEEGACCKDRLMAGDQILEVNGIDLKRAETHDEAINVROTPQVR 640
 DB 1786 GADTVLGVIHVEVSDGAHAGRLKPDQVLENGSLKRVTHDOSIATVIRTPPVYR 1845
 GADTVLGVIHVEVSDGAHAGRLKPDQVLENGSLKRVTHDOSIATVIRTPPVYR 1845
 DB 641 LTLVRDE-----APYKEEVEDTLTLELQKPKGGLSTVGRNDTGVFVSDIVVGIAD 696
 LTLVRDE-----APYKEEVEDTLTLELQKPKGGLSTVGRNDTGVFVSDIVVGIAD 696
 DB 1846 LLIYDVNLQSLDLPQTQVNFELDVKKTRGIGISIVGKNRPVYVSVIVAGLAE 1905
 LLIYDVNLQSLDLPQTQVNFELDVKKTRGIGISIVGKNRPVYVSVIVAGLAE 1905
 DB 697 PDGRLIOGDIILVNGEDVRNASSQAVALLKCSIGTYTLEVGRIKACPF----- 746
 PDGRLIOGDIILVNGEDVRNASSQAVALLKCSIGTYTLEVGRIKACPF----- 746
 DB 1906 SDGRIMTGDQILEVNGKRCQMDVAMLTKTIGKYLKTEENNNDPVPRVAVAVSP 1965
 SDGRIMTGDQILEVNGKRCQMDVAMLTKTIGKYLKTEENNNDPVPRVAVAVSP 1965
 DB 747 ----- 746
 ----- 746
 DB 1966 AATPSTASVYTSAPNVSPEPNVVEPIPGTASAREENDVPPAPMRPIITHTSPE 2025
 AATPSTASVYTSAPNVSPEPNVVEPIPGTASAREENDVPPAPMRPIITHTSPE 2025
 DB 747 --HSERRSPQTSQVSEGLSFTF-----PLSGSTSESLSSSKNALASEI----- 792
 HSERRSPQTSQVSEGLSFTF-----PLSGSTSESLSSSKNALASEI----- 792
 DB 2026 GCEIQDEPAGLSPTVEPSSGCDPFSVQEEBPRTPTTTT-----SSNNNNSLAIDIIIDL 2082
 GCEIQDEPAGLSPTVEPSSGCDPFSVQEEBPRTPTTTT-----SSNNNNSLAIDIIIDL 2082
 DB 793 --QGLRT--VENKKGPRTSLGISIAGVCSPLGDVPIFIAMHPGVAQAQOKLAVGDR 848
 QGLRT--VENKKGPRTSLGISIAGVCSPLGDVPIFIAMHPGVAQAQOKLAVGDR 848
 DB 2083 KEESGDTLLVELKRYVDQIGM-----GIGK--RSRGILVTSIQPSAA--EKLKVGDR 2134
 KEESGDTLLVELKRYVDQIGM-----GIGK--RSRGILVTSIQPSAA--EKLKVGDR 2134
 DB 849 VTICGSTEGMHTQAVNLKNAAGSIEMQV 879
 VTICGSTEGMHTQAVNLKNAAGSIEMQV 879
 DB 2135 IAV--NALPVSQDLSAVTFVKASGORLYDQI 2163
 IAV--NALPVSQDLSAVTFVKASGORLYDQI 2163

RESULT 4
 T09457
 numb-binding protein LNX80 - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 02-Sep-2000
 C:Accession: T09457
 R:Bio. S.E.; Jacob, S.; Wolting, C.D.; French, M.B.; Rohrschneider, L.R.; McGlade, C.
 A:Title: The mammalian Numb phosphotyrosine binding domain: Characterization of bindi
 A:Reference number: Z16678; MUID:98204916
 A:Accession: T09457
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA

A:Residues: 1-728 <DHO>
A:Cross-references: EMBL:AF034745; NID:93041878; PIRN:AAC40075.1; PBD:93041879
A:Experimental source: developmental stage: embryo
C:Genetics:
A:Gene: lux
A:Superfamily: RING finger homology
F:41-88/Domain: RING finger homology <RRN>

Query Match	11.6%;	Score 589;	DB 2;	Length 728;
Best Local Similarity	29.1%;	Pred No. 5.4e-26;		
Matches 188;	Conservative 115;	Mismatches 238;	Indels 106;	Gaps 23;

```

0Y 411 SEFKVVOHLELPD-----QGGL--GIAISEEDLYSGVITKSTLEHCAVTOADRLKVGQD 463
0Y 130 TSCKASHGSLKDRKRNRDGDGPOGCASLMAITLS-----PVSAMAT----- 173
0Y 464 IIAVDEIVYGPRIEKFISLTKTAKTVKLTTHAENPDQOAVP---SAGASAEGKKNS 520
Db 174 ISLMDDEPLDNP--AYVSVSEDEGEFVANSSDGSRNRTARPFERSYMSSEFKKINRA 231
0Y 521 OS-LWVPOGSEPEESI-----RMTSRSPAIAPASPPACPIIPGE-TTIEISKGTGL 574
Db 232 LSAIARTKSGSVVANNVDGGRDNSENVTYPEVE---PRLFHLIPDGEIISIKINRADPSE 268
0Y 575 GJST--VGSDBLLCAFTTHEYEEGACCKDGRYLAAGDQILEYNGIDLKRTATDEAIIVL 632
Db 289 SLSIRLVGSEPEPLVNIITQHTYRDGVIAARDGLTGLDIIILKYNMGDISVNPYNAVRL 348
0Y 633 ROTPORVRLTYLRD-----APYKEEVEYDUTLTIELOK-KPGKGLSLIVGKRNDT 682
Db 349 ROPCVLRLTYLREQKFRSRNANHVDSYGRPDSDHVLITKSSPEOGLIKTLRVDER 408
0Y 683 GVEFSDIYVGGIADPRGRITOGDQILVNGEDVRNANSOEAVALLKCSGTVTLVEGRK 742
Db 409 GVFIFNVNLGVADRIQGOLEENDRVLAINGHDLRESPESSAHILIASERKVALVYSR-- 466
0Y 743 AGPFHSERRPS-----QTSVSEBGLSSTFFPLSGSSTSELSSESSKKNALASEIOGLRT 757
Db 467 -----QVROSSPDIDIOEAGWISNGOOS---RPPGR-----NTASKRAPATCHE---KV 508
0Y 758 VEMKKGPTSLISISINGVSGSPLDGVPRTFIAMNHPGVAAOQTKLRAGDRIYVTCGSTE 857
Db 509 VSVMKDPSDSLMTYGGGASHREMDRLPIYIVISVEPGVYSRDBCRITGTDLILVNGIELT 568
0Y 858 GHTHQAVVLLKNNAAGSILEMOYVAGGDVSVUTGHNODRPASSLSSTFGLTSTSTFODDLCP 917
Db 569 EVSRTAEVAAILKSAPBSVYLKAL---EYK-----BOEAQEDCSPALDS---NNHVTP 615
0Y 918 P-----QCKSTYLERGPDG-LGFSIYGVGSGSPHAGDRLTYVYKFAAG 958
Db 616 PCDWSPSWMMTELDPQYLNCNKDVLIRNTAGSLGICGIVGYEYSGNKPRFIKSLIVEGT 675
0Y 959 AASEDGRILKRGQIIITAVNGOSLEGVTHEAVVAILKTKGTVTPLMVL 1005
Db 676 PAVNGRILKRGQIILAVNGRSTSGMTHACLARKLTKLKRITVLTJLAS 722

```

```

RESULT      5
T09458
numb-binding protein LNXp70 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T09458
R:Dho, S.E.; Jacob, S.; Wollring, C.D.; French, M.B.; Rohrschneider, L.R.; McGlade, C.J.
J. Biol. Chem. 273, 9179-9187, 1998
A:Title: The mammalian Numb phosphotyrosine binding domain: Characterization of binding
A:Reference number: Z16678; MUID: 98204916
A:Accession: T09458
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Residues: 1-628 <DHO>
A:Residues type: mRNA
A:Cross-references: EMBL:AF034746; NID: g3041880; PIDN: AAC40076.1; PID: g3041881

```

A;Experimental source: brain
C;Genetics:
A;Gene: 1nx

Query Match	11.5%	Score 586;	DB 2;	length 628;
Best Local Similarity	29.0%	Pred. No. 6.4e-26;		
Matches 187;	Conservative 114;	Mismatches 237;	Indels 106;	Gaps 23;

```

OY 414 KVVQHLLEPKD-----QGL-GLISEEDTSLGVIKSLFTHGVAADGRKLKVDQJLA 466
Db 33 KGASHYGLTKOKRRKSQDCPCDCAISMAITLIS-----PEVSAAT-----ISL 76
OY 467 VDEIVGVPIERKFTLSLKTAKMTVKLTTHAENPDQAVP---SAAGAAGEKKNSQS- 522
Db 77 MDEGEGLDNP--AYVSVEDGEPVANSSDGRKSNRTPARPFERSJMRSHRSFKINRALSA 134
OY 523 LMVPOGSGPEPEST-----RNTSRSTPAIFASDPATCTPIBCE-TTIEISKGRFGLIS 577
Db 135 LKRTSGSGVANHVOGRDNSENTEIVPEVF---PRLPHLIPGELTISIKINADPESEIS 191
OY 578 I--VEGSDTLGCAFTIHEEYEEGACCKGRRLAGAOIIEVNCIDIRKATTHDEAINVLROT 635
Db 192 IRLVGSETEPLVHTIIOHTYRDRGVARABGRLLPGIILKVMGMDISNPNHNAVALLQAP 251
OY 636 PORVRLTYLDE-----APYKEEVCDTLTTELQK-KPKGLGSLVGRNDQVF 685
Db 252 COVLRLTLYLREQKFRSRNANHPDGYGPRDGSFHYILKNSSPEEDLGILKVLARVDEPVEF 311
OY 686 VSDIYKGGIADPDGDLIGODOLILVNGEDVNRASOEVAALLKCSLGVTVEVGRKMG 745
Db 312 IFNVANGVADRKHGLLENDKRYLAINGHDLRGRGSPESANHLIQASERYVHLVSR----- 366
OY 746 FHSERRPS-----QTSQVSEGLSSEFTPEPLGSGSTSESLESSKKNALASEIQLRTYEM 800
Db 367 --QVROSSPDIFQEGAGWISNGQS---PGPER-----NFRASKPAATCHE---KVYSV 411
OY 801 KKGPDLSIGISAGVGSPPLGDVPRFIIMMHPHTQVAAQOTKRVYDRIYTCGTSEBMT 860
Db 412 WMDPESJLGMVYGGASHREMDPLTYIVISVEGGVYISKRGKTKTDIILNNGILELTVS 471
OY 861 HTOAVNLKKNASGSIEMQVVAAGDVSVYTGHHQEPASSLSLFTGLTSTISIFODDLGAP- 918
Db 472 RTEAVAILKSAKPSVYLKAL---EVK-----EQDAQEDCSPALDS---NHNTTPPGD 518
OY 919 -----QCKSTILEGPRGC-LGFSIVGCGSPHGDLPLYVTVFKAQAS 961
Db 519 WSPSWMMWLELPQYLNCNKRDVILRRNTAGSLGFCFVGGYEESGNNKPFIKISIVGCTPAY 578
OY 962 EDGRLLKRGDOJIAVNGOSLEGVTHEEVAAILKRTGVTYLTWMLS 1005
Db 579 NDGRIRCGDILAVNGRSTSGMITHCLARMLKELKGRITLTLTAS 622

```

```

RESULT      6
T233160
hypothetical protein K01A6.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Aug-2000
C:Accession: T233160
R:Cottage, A.
submitted to the EMBL Data Library, January 1996
A:Reference number: Z19701
A:Accession: T233160
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-1012 <M12>
A:Cross-references: EMBL:Z68750; PIDN:CAA92965.1; GSPDB:GN00022; CESP:K01A6.1
A:Experimental source: clone K01A6
C:Genetics:
A:Gene: CESP:K01A6.1
A:Map position: 4
A:Intons: 107/1; 151/2; 183/2; 270/2; 316/3; 410/3; 628/3; 922/3; 971/3

```

C:Superfamily: WW repeat homology
F:131-168/Domain: WW repeat homology <WWR>

Query Match 10.5%; Score 534; DB 2; Length 1012;
Best Local Similarity 24.1%; Pred. No. 1.3e-22;
Matches 208; Conservative 123; Mismatches 298; Indels 224; Gaps 31;

QY 204 NGLTRGRIVAPPSQSEPEKAPLCSP-----PPPSAFAMGSDHTQSSAKIS 256
DB 317 SCLTL-----SSSSPPDSGFSSPTRRKFGDPPEATTSADYDHSMKFSRS 365
QY 257 QVDEKEDGYSWKNIENRYGTGLHWEIEKGSGGLSLAGKDKSRMSVFI--VG 314
DB 366 NLFETDP-----ARLGGELISTKIYKCAKGLGFTLIGDSSSRKDEFTIQKS 413
QY 315 IDPNAGAKDRLQIADLELEINGQILYGRSHQNASIICAPSKVKIIFIRNKDAVQ- 373
DB 414 VLSGPPAANGVLRSGDILVRNGRLLGATQKEACDVVAIP-----VNEA 460
QY 374 --MAVCPGNAVEPLPSNSNLONKEPEVTYTSDAVDSLSEKKNQHLPRDQGLGIA 431
DB 461 VDIQVCRGELFIDPAN-----RIVENYAAAKS-RDLHEIDIFKSGEGFET 508
QY 432 ISEEDTLGVIITKSTHEHVAATDRKLVGDQILAVDEIVVGYPIEKFIISLTKAKTV 491
DB 509 IA--DNINQRIKKIL--FPSCPNLMGCDITVELDGNVAPRPHQVLMLEKREPICT 563
QY 492 --KLTIHAENPDQVAPSAAGAAGEKKNSSQSLMVP-----QSGSPESIR 537
DB 564 RCKLVVKKRSPKTRSTPFAAFRGEPOJNMDSAPLVRSKTPAERQTSREEDQNR 623
QY 538 NRSRSTAFIFASDPATCIIIGCETTEISKGRGLGISTYGSPTLLGATIIHEVYE 597
DB 624 NT-----LORQPAVYSEMGMSAIPASRRP-----SSTLGF----- 657
QY 598 GAACKDRLMAGDQILEVNGIDLKRAKHDEAINVLRQTPQRYRLTYRDEAPYKEEVED 657
DB 658 -----ATPNYIPISQY-----NQKPSD 674
QY 658 TLTIELQKPKGGLGISTY--GKRNDTGVFVSDIVKGIADPDGRLIQDQILLVNGEDVR 716
DB 675 LITVSLIRKP-VGFGFRLGIVESKPTLSVGOIVIGAAEEDGRLOEGDEIVEIDGHANE 733
QY 717 NASQEVALLKCSLGT--VTELVGRIRKGFHSERRPSTQOVSGSISSTFPPLSSG 774
DB 734 GASHSEAVVLLDAONKHKVLLI-----RRERSTPARRKSLNS--AGPS 777
QY 775 TSESSLESSSKKNALASEIOGLRTVEKKKGPDTSLGISIAGV---GSPLDVPIFIAMH 831
DB 778 -----GSYDVLHNRNENGFGFVLMSSQHKNGSTVGOI-----Q 811
QY 832 PRGVAAQIOKLVAGDRIVITIGCTSTEGKTHQAVNLKNAAGSIEMQV---VAGDSV 887
DB 812 PPSPAKRCRLSVGRVIVANGIDILSLHPDTISLIKSGLSVRLTIAPRYAGVFLPM 871
QY 888 VY-----GHHQ-----EPASSLSFTGLTISIRIQDGLG--PRQCK 921
DB 872 VATATLGRNTMGHTESNGLPRPPPSVEKHPPSYLLFDGLS---INDRSMNGNLI 927
QY 922 STLEKPGDLGFSIYGVGSPHGLPIYVKTVFPAKGAASEDGRKLRQDIIAIVNGOSLE 981
DB 928 DVTLEKGTGFGFSIRG--GQFEGSMPLFLVLIADGPAKADGRLOYGQTLTINQSTK 985
QY 982 GVTHEEAVAILKRTKGTVTLMVL 1004
DB 986 GMSHDAIRIKQ-HTMVLVL 1007

RESULT 7
571625
protein-tyrosine-phosphatase (EC 3.1.3.48) RIP - mouse
N/Alternate names: epidermal growth factor-binding protein; serine proteinase

C:Species: Mus musculus (house mouse)
C:Date: 27-Nov-1997 #sequence_revision 12-Dec-1997 #text_change 20-Jun-2000
C:Accession: S71625; S67987
R:Chida, D.; Kume, T.; Mukoyama, Y.; Tabata, S.; Nomura, N.; Thomas, M.L.; Watanabe, F.B.S. Lett. 358, 233-239, 1995
A:Title: Characterization of a protein tyrosine phosphatase (RIP) expressed at a very
A:Reference number: S71625; MUID:95145716
A:Accession: S71625

A:Molecule type: mRNA
A:Residues: 1-2450 <CHI>
A:Cross-references: EXBL.D83966; NID:g1232103; PIDN:BA12158.1; PID:g1232104
A:Experimental source: strain DBA/2; cell line MEL 745A
R:Wolf, B.B.; Brown, M.D.
FEBS Lett. 376, 177-180, 1995
A:Title: Epidermal growth factor-binding protein activates soluble and receptor-bound
A:Reference number: S67987; MUID:96105375
A:Accession: S67987
A:Molecule type: protein
A:Residues: 1098-1102 <WOL>
A:Experimental source: submaxillary glands

C:Genetics:
A:Gene: Ptpn13
A:Map position: 5
C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 13; GLGF domain homolog
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; transmembrane protein; cy

F:566-860/Domain: protein 4.1 membrane-binding domain homology <B41>
F:1089-1165/Domain: GLGF domain homology <GLG1>
F:1361-1437/Domain: GLGF domain homology <GLG2>
F:1495-1574/Domain: GLGF domain homology <GLG3>
F:1769-1840/Domain: GLGF domain homology <GLG4>
F:1863-1937/Domain: GLGF domain homology <GLG5>
F:2203-2422/Domain: protein-tyrosine-phosphatase homology <PPT1>
F:2374/Active site: Cys (phosphocysteine intermediate) #status predicted
F:2380/Binding site: substrate phosphate (Arg) #status predicted

Query Match 10.3%; Score 523; DB 2; Length 2450;
Best Local Similarity 23.7%; Pred. No. 2.1e-21;
Matches 240; Conservative 134; Mismatches 346; Indels 294; Gaps 32;

QY 196 LEDSPAGKNGTLKPGDRIVEAPDSQSEPEKAPLCVPPPPSAPFAEMGSDHTQSSAKI 255
DB 1012 LNSKSVASLNSPERRNIESDSSTE-DPGQAVYVGMSP-----SSGKS 1055
QY 256 SODVDEKEDFGYSWKNIENRYGTGLG--ELMIIELEK-GHGLGSLAGNDRSR--MS 309
DB 1056 SSQVPPKDN-----DTLHRKMSIVSSPERETLVNLKPKKRGGLGQIIGKMGKGLDG 1110
QY 310 VPIVIGIDPAGAGKDRQIADLELEINGQILYGRSHQNASIICAPSKVKIIFIRNKD 369
DB 1111 VPIASVTPGPRADLDGCLKPGDRLLSVNSVLEGVSHAAVQIILQNAPEADYLVISQPRE 1170
QY 370 AVNQAAVCPGNAVEPLPSNSNLONKEPEVTYTSDAVDSLSEKKNVQHLPRDQGLG 429
DB 1171 KPSK-----VPSYVHFAMGMSYTKPKPAYMDSADPSE-----DQPMFRTGLR 1215
QY 430 IAISEEDTLGVIITKSTHEHVAATDRKLVGDQILAVDEIVVGYPIEKFIISL----- 484
DB 1216 HTPESPFGSLG---GLRGSSLSQDSR---TESASLSQSOVNGF---FASHLGRDGM 1263
QY 485 -----RTAKMTVKLTIIHAENPDQVAPSAAG-----AASGEKNSS-QSLMVPQ 527
DB 1264 QEPQHSPPSPVTTKYNETFSDSNRSAKKRGISDLIHLDCABSDKRDSTYTSQDHQ 1323
QY 528 SGPSPESIRNTS-RSSTPAIFASDPATCIIIGCETTEISKGRGLGSLIVGSDTL 586
DB 1324 TSKQEPSSSLSTSNKTSFPTSSASPK-----PGDTFVELAKTKGSLGISTVGAVNSV 1378
QY 587 --GAFIIEVEEGAACKDRLMAGDQILEVNGIDLKRAKHDEAINVLRQTPQRYRLTY 644
DB 1379 RHGGIYVKAIIIPKGAASEDGRIVKGRVLANGVSLGEGATHRQAVETLANTQOVHLLLE 1438
QY 645 RDEAPYKEE-----VCDTLTIELQK-KPKK 669

```

Db 1439 KGQVPTSHEDPAGPQSPPPDQDAORQAPKAVAKOTPHVKDYSPVTEEDNTEVKLFKNSS 1498
QY 670 GLGLS-----YGRKNDTVGVPSDIYKGIADPDGLIGDDOILLVNGEDVRNASEA 722
Db 1499 GLGFSFSEHDNLIEQDINGSIVRVKLFEGGPAESKIDVDYLKNGAPLGLSQDD 1558
QY 723 VAALLKSGISGVTELEVGRKAGP-----FHSRRPSQTSQVSEGL 763
Db 1559 VISALRGAPVPSLLCPCAPGVLEIDTFLNPLYSFANFLMSKRTSQSPSSVEGA 1618
QY 764 SFFPPPLSG-----STSES----- 778
Db 1619 SSHNGVSGKTKNHCRAFPSRESYSDSHSEGEDSVAPAKMNVTRVAAPRPAESQE 1678
QY 779 -----LESSK----- 784
Db 1679 ESICAMFLPKITGKLESESHPPPLDVSFGQTCOPPAECAPSDATGKHETHLASOLSK 1738
QY 785 -----KNALA-----SEIQGLRTVEKKGPTDLSGISIAG--VSPPLGDVPFIAM 829
Db 1739 ENITTLKNDGNHLEDESELEVLLITLVKSEKSGISGTVTKGSQSTIGCYVHDV-----I 1793
QY 830 MHPGVAAQOTKLVGRIVITICSTEGMTHQAVNLKNASGSIEMQVAVAGDVSYYT 889
Db 1794 QDP---AKGDRRLKAGRLKIVNDVTYNTMTDAVNLLRAAPKTVRL-----VL 1840
QY 890 GHNOEPASSLSFTGLTSTSFODDLGPPCKSTILEKPGDGFSGYSIGVSGPHGLPI 949
Db 1841 GRILE-----LPRMPYF-----PHLLPDITVTCHEGSELGFSLSGGSGPHG--VV 1883
QY 950 YKTVFPAKGAASEDGRKRGDQIIIAVNGQSEGVTHEEVAAILKRTNGTVTLNV 1003
Db 1884 YISINPRSAAVADGSLDLDIIHYVNGVSTQGMLLEDANRALDLSPLSVLKV 1937

RESULT 8
167629
protein tyrosine phosphatase (PTP-BAS, type 2) - human
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 16-Jul-1999
C:Accession: I67629
R:Meekawa, K.; Imagawa, N.; Negamatsu, M.; Harada, S.
FEBS Lett. 337, 200-206, 1994
A:Title: Molecular cloning of a novel protein-tyrosine phosphatase containing a membrane
A:Reference number: I53483; MUID:94116679
A:Accession: I67629
A>Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Residues: 1-2466 <RES>
A:Cross-references: GB:D21210; NID:9452191; PIDN:BA04751.1; PID:9452192
C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 13; GLGF domain homology:
F:574-668/Domain: protein 4.1 membrane-binding domain homology <B41>
F:1354-1430/Domain: GLGF domain homology <GLGF2>
F:2218-2437/Domain: protein-tyrosine-phosphatase homology <PTP>

```

```

Query Match 10.0%; Score 509.5; DB 2; Length 2466;
Best Local Similarity 22.3%; Pred. No. 1.3e-20;
Matches 256; Conservative 146; Mismatches 395; Indels 351; Gaps 41;

```

```

QY 142 ELQNTAVSNMQRPRVELMREPKSLGISYGGMGSRLSNG----- 184
Db 868 QLOMRAROSNODADIERASFRLNLQAESVRFNMGRALSTGSLASTLKLAVRPLSV 927
QY 185 --EVMRGIFIKHY-----LEDSPPAKNGTL-----KPGD--RIVAPRSESEPEKAPLCS 231
Db 928 QAEILKRIKSCSELVLYPLQNSKEKNDKASWEKPRMKSSTYHDLQASLTPRKKNIV 987
QY 232 VPPPPSAFAEM-GSDHTOSSAS-----KISQVDKED 263
Db 988 NMERPPQTVAVELVGKPSHONSRSDAESLAVTKLNNSKSVASLNRSPERRKHEDSSSIE 1047

```

```

QY 264 EFG--YSWKNIERYGTLTG---ELHMLELEK-GHSGLGSLSLAGNKRDR--MSGVYGI 315
Db 1048 DPGAGVLDVLRKMSIYSSPEREITLVNLLKKAQKGLGFGPIIGEKMKRGLDIFISSV 1107
QY 316 DPNAGKGRGLQIADLELLEINGQILYGRSHQNASIHKCAPSVKIIIFRNKAVNOMA 375
Db 1108 APGCPADIDGCLKPGDRILISVNSVSLGESHHAILEIQNAPEDVTLVISOPEKISKVP 1167
QY 376 VCPGNVAPLPSNSENQNKTEPEVTTSDAAVDLSEFKN-----VOHLELP--DOGG 427
Db 1168 STPVH---LTNEKNTWKSS---YMODAIDSSKDHMSGTLRHISNSFFGSGG 1219
QY 428 L'-GIAISEEDTLGVILKSLTEHGVAAVDGRKLVQDQIILAVDEIYVGYPIEKFIILK 485
Db 1220 LREGSLSDQDRETSASLSQVNGFPAS---HLGQGTW---QESQHGSPSPVIS--- 1269
QY 486 TAKMTVKLTIAENPDQANVPSAAGASGEK-----NSQSILMPQSGSPPEPES 535
Db 1270 --KATEKETFTDSNQSTKPKGISDVTDYSDRGSDSDEATYSSSDHQTQPKQ---ESSS 1324
QY 536 IRRNRSSTPAIFASDPATCPPIPGCEETIEISKRTGLGISYIGSGDTLL--GAFIIE 593
Db 1325 SVNTSNKMNFTFESSPEK---PGDIFEVELANDNSLGIYTGAVNTSVRHGGIYKA 1380
QY 594 YVEBGAACKDGRILWAGDQIILEVNCIDLRKATHDAIVNLROTPOVRRLTYRDEAPYKE 653
Db 1381 VIPQGAESDGRIRHKGRVILAVNGVSLGATHKQAVETLRNTGQVHLLLEKQSPFSKE 1440
QY 654 EV---C-----DPLTILELCKPKKGLGLS----- 674
Db 1441 HVPVTPQCTLSDNAQAGPEKVKTKTQVKDYSFVTEENFEVLEFKN--SSGLGFSFSRE 1499
QY 675 ---IVGRNDTVGVPSDIYKGIADPDRLQDQIILVNGEDVRNASEAVALLCSL 731
Db 1500 DNLIPEQINASTVAVKTLPEQPAAESGKIDYGVILKVNASLKGISQGEVIALMGTA 1559
QY 732 GTVTELEVGRKAGP-----PHSE----- 749
Db 1560 PEVFLLCRPPPGVLEIDTALLTPLOSPAVOLPNSKSDSQSPCVQSTSDENEMSK 1619
QY 750 ---RRPSQTSQVSEGS-----LSSEFT----- 770
Db 1620 SKKCKCKSPSRDSTSDSGSGEDDLVTAPANISNISTWSALHQTLSMNVSOAQSHEAPK 1679
QY 771 SGSST-----SESSLESSK----- 784
Db 1680 SQEDTCTMFFYPPQKINKPEFEDSNPSLPPDMAPOGVSQPOSEASSSSMOKYHITHI 1739
QY 785 -----KNALASEIQGLR-----IYEMKKGPTDLSGISIAG--VSPPLGDVP 824
Db 1740 SEPTROENMTPLKNDLNNHLEDFELELLITLKSEKSGISGFTVTKGNORIGCYVHDV- 1798
QY 825 IFIMMHPGTVAQOTKLVGRIVITICSTEGMTHQAVNLKNASGSIEMQVAVAGD 884
Db 1799 ----IQDP---AKSDGRILKPEDRLIKVNDVTYNTMTHTDAVNLLRAASKTVRL----- 1844
QY 885 VSVVTGHHOE-PASSLSFTGLTSTSFODDLGPPCKSTILEKPGDGFSGYSIGVGYSP 943
Db 1845 ---VIGRVLELPRPML-----PHLLPDITLVTCNKRELGFSLGCGHDSL 1885
QY 944 HGDLPYVKTVFPAKGAASEDGRKRGDQIIIAVNGQSEGVTHEEVAAILK 994
Db 1886 Y--OVVYISINPRSAVAILGQDLDDIVHYVNGVSTQGMLLEVRNARALDMSPLSVLKA 1943
QY 995 TKGTVTLM 1002
Db 1944 TRNDLPVY 1951

```

```

RESULT 9
167630
protein tyrosine phosphatase (PTP-BAS, type 3) - human
C:Species: Homo sapiens (man)

```


A:Cross-references: GDB:306348; OMIM:600267
 A:Map position: 4q21.3-4q21.3
 C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 13; GLGF domain homology;
 C:Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; tyrosi
 F:574-868/Domain: protein 4.1 membrane-binding domain homology <B41>
 F:1089-1175/Domain: GLGF domain homology <GLG1>
 F:1373-1454/Domain: GLGF domain homology <GLG2>
 F:1511-1590/Domain: GLGF domain homology <GLG3>
 F:1799-1870/Domain: GLGF domain homology <GLG4>
 F:1893-1967/Domain: GLGF domain homology <GLG5>
 F:2242-2461/Domain: protein-tyrosine-phosphatase homology <PPT1>
 F:2413/Active site: Cys (phosphocysteine intermediate) #status predicted
 F:2419/Binding site: substrate phosphate (Arg) #status predicted

Query Match 9.7% Score 494.5; DB 1; Length 2490;
 Best Local Similarity 22.6% Pred. No. 9.2e-20;
 Matches 243; Conservative 140; Mismatches 371; Indels 319; Gaps 38;

QY 162 EPRKSLGISIVGRCMSRLSNGEVMRIGITKHVLEDSPPAKNGTLKPGDRIVEAPSOSE 221
 DB 990 EPPQTVAEVLGKPSHQRSDAELAGV---TKLNNSKVASLNRSPERRKHSSSSSI 1046
 QY 222 SEPEKAPLCVPPPPPSFAEMGSDHRTSSASKISQDVDEDEKYSKKNIREKIGITLG 281
 DB 1047 EDPGQAVYLGMT-----MHSSGNSSQVPLKENDV-----LHKRWISVS 1086
 QY 282 ---ELHMLEK-GHSGGLSLAGNKRDR--MSVFIQVDPNGAAGKGRLOADELLE 335
 DB 1087 PERETVLNKKDAKYGFGFDIGEEKMGRDLGIFISSVAPGAPADLGLCKPDRILIS 1146
 QY 336 INGQLYRSHQNASIIKCAKPSKYLIFIRNKDAVNOMAVCPGNAPVPLPSNENLONK 395
 DB 1147 VNSVSLGVSHHAAIEIIONPEDVTLSQPKKIKVSPSTPVH---LNNEMKNMYKK 1202
 QY 396 ETEPTVTSDAVDSLFKN-----VOHLELPK--DQGL--GTAISEPTLSGVITKS 445
 DB 1203 SS---YMODSAMSSSKDHMSKRTLHISENSRPGSGGLREGSLSDSTESASISQ 1258
 QY 446 LTHGVNATDRLKLVGDQILAVDEIVVGVPIEKIFSLTKAKMTVKTLIAENPDQAV 505
 DB 1259 SQVNGFFAS---HLGDQTV---QESQHGSPSPVIS-----KATEKETFDSDNSKTKK 1306
 QY 506 PSAGAASGEKK-----NSSQSLMVPQSGSPPESTIRNTRSRSTPAIIPASDPATC 555
 DB 1307 PGISDVTYSDRGSDMDDEATYSSQDHQTPKQ--ESSSVNTSNKKNFKTFSSPPK- 1362
 QY 556 PIRPCETTIEISKRTGLGSIY-----GGSDPLL--GAFIHEVYEGGAACKDGRIMA 608
 DB 1363 ---FODIREVELAKNDNLGISTVYVLPDKGVNTSVRHGCIYKAVIPQGAESDGRILK 1419
 QY 609 GDOILEVNGIDLKRAIHDEAINVLKQTPQVRVRLTLYRDEAFYKEEV-----C----- 656
 DB 1420 GDRVLAVNGVSELEGATHKQAVETLRNTQGVVHLLLEKQSPTSKEHVVPVPOCTLSQNA 1479
 QY 657 -----DTLTLEOKKRGKGLGSL-----YKRRNDTGCVFV 686
 DB 1480 QGQGEPEKVKTTQVYKSEVTEENTFEYKLEKN--SSGLGFSESRDNLIPQINASIRV 1538
 QY 687 SDIYKGIADPDRLIOGDOILLVNGEDVYRNASQAVALLKCSIGVTLTEVGRIRKAG-- 744
 DB 1539 KKLFPQGAASGKIDGVILKYNGLSKGLSQOEVLSALRGTAPEVFLLCRPPCVL 1598
 QY 745 -----PRHSE-----RRSOTSQVS 759
 DB 1599 PEIDTALLFPQSPAQVLPNSSKDSQSPCEQSTSSDENEMSKKQCKSPSRSDYS 1658
 QY 760 EGS-----ISSFTF-----PLSGST----- 775
 DB 1659 DSSGSGEDDLVTAPANISNSTWSALHQTLSNMVSOASHHEAPKSOEDTICTWFFYPOK 1718
 QY 776 -----SESLSSSK-----KNA 787

DB 1719 IPNKPEFEDSNPSPLPDPMABQSQYQPOSASSSSSMDKYHIHISEPTROENNTPLKND 1778
 QY 788 LASEIOLGR-----TYEKKKQPTSLGISING--VSGPAGDVPFIAMHPTGVAQT 839
 DB 1779 LKNHDEFELEVELLITLIKESKSLGFTYKGMORIGCYHVD-----IODP--AKSD 1830
 QY 840 QKLRYGRIYVTCSTEGMTHTQAVNLKNASGSIEMQVAGGDVSVYTGHNQ--PASS 898
 DB 1831 GLRKPGRDLIVNPDVNTNMHTDAVNLRAASKTVRL-----VIGRVLELPRIIP 1880
 QY 899 SLSTGLTSTIFODDLGPPCKSTTLERPGDGLGFSYGVGYSBPHGLPIYKTVAKG 958
 DB 1881 ML-----PHLLPDTITCNKEELGFFSLCGHSDSLY--QVYIYSDINPRS 1922
 QY 959 AASEDRGLKRGDOIIAVNGSLGVTHEA-----VALKRTGTATIM 1002
 DB 1923 VAAIEGNQLQDLVHYVNGVSTQGMTLEEVNRALDMSLPVLAKATKNDLPV 1975

RESULT 11

JE0209
 brain-specific angiogenesis inhibitor-associated protein 1 - human
 N:Alternate names: BAI1-associated protein 1; BAP1 [misanomer]
 C:Species: Homo sapiens (man)
 C>Date: 21-Aug-1998 #sequence, revision 21-Aug-1998 #text, change 18-Aug-2000
 C:Accession: JE0209
 R:Shiatsuuchi, T.; Futamura, M.; Oda, K.; Nishimori, H.; Nakamura, Y.; Tokino, T.
 Biochem. Biophys. Res. Commun. 247, 597-604, 1998
 A:Title: Cloning and characterization of BAI1-associated protein 1: A Pdz domain-conta
 A:Reference number: JE0209; MUID:98321173
 A:Accession: JE0209
 A:Molecule type: mRNA
 A:Residues: 1-1256 <SH1>
 A:Cross-references: GB:AB010894; MID:93370997
 A:Experimental source: brain
 C:Genetics:
 A:Gene: GDB:BAIAP1; BAP1
 A:Cross-references: GDB:9864783
 A:Map position: 3p14.1-3p14.1
 C:Superfamily: WW repeat homology
 F:300-337/Domain: WW repeat homology <WW1>
 F:359-396/Domain: WW repeat homology <WW2>

Query Match 8.9% Score 453.5; DB 2; Length 1256;
 Best Local Similarity 22.4% Pred. No. 7.1e-18;
 Matches 184; Conservative 126; Mismatches 283; Indels 229; Gaps 30;

QY 378 PGNAVEPLPSNENLONKETEPTVTTSDAANDLSSFKVQVHLELPKDOGGIGLAI---SE 434
 DB 442 PSN---PEPAPEVPLQGR---PFTTRNPSELK---GKFIHTKLKSSRGFGFTVVGDE 491
 QY 435 EDTLSGVTKSLTETHGAATGRLKVGDOILAVDEIVGVPIKFISSLTKATMTVK-- 492
 DB 492 PDEF--LQIKSLVLDGRALDGMKMETGDVYVMDTCVGLGTHAQVKIRFSITIGASVD 549
 QY 493 -----LTIHAENPDQAVPSAAGAASGEK-----NSSQSLMVPQSGSPPEPSIR--N 538
 DB 550 LELCRGYPLPDPDPDPTSLVTSVAIL--DKEPLIYNGCETYSPPASHSKTKGVNGMK 606
 QY 539 TSRSTPAIFSDP-----ATCPITPGCETTIEISKRTGLGSLY---G 580
 DB 607 DARPSSPADVASNSHGVPNDTVLASSIAQPEL---IIVHIVKGMGFGFTIADSPG 662
 QY 581 GSDTLLGAFIHEVYEGACCKDGRLMNAGDQILEVNGIDLKRAIHDEAINVLROTPOVR 640
 DB 663 GGGQGVKQIV-----DSPRCRG--LKEGLDLYEVNKKKNQOALITHNOVDMVLCECPGSE 714
 QY 641 LTL-----YRD-----EAYKE 652
 DB 715 VTLLVQNGGLVPKPKSPQLEKKDSQNSQSHSVSHRSRLHTASPSHSTQVLEPFPAL 774
 QY 653 EEVCDTLTLEOKKPG----- 668

Db 775 AQPOTDSSGGKRPDPFKIWAQSRMYENRPMSPASGLSKGERERINSTNGECI 834
 QY 669 -----KGLSTIVKRNNG--VFVSDIYKGIADPDGRLIGDQILLVNG 712
 Db 835 PDYQODIFLWKEKKEGFGFRILG-GNEPDEPIYGHIVPLGADPDGRSDDELICVDG 893
 QY 713 EDVRNASEAVALLK--CSLCTVTLVGR-----IKAGFHSERP--S 753
 Db 894 TPVICKSHQVLMQAAKQKHVNLVARKVFAVPKTEENVPASSHSSNPASLT 953
 QY 754 QTSQVSEGLSFTPLPSLSTSELES--SKKNALASEIQLRTYEMKKGPTDSLGISI 812
 Db 954 EKKRPQGSQNLNTVSSSGSTSGISGGSGSVIVVQPYVEIRRGNEGFGFI 1013
 QY 813 AGVGSPLGDVPI---FLAMHPTG-----VAAQTKLRVDRIVTICGTSTEGMHT 862
 Db 1014 VSSVRPEAGTTFAGNACVAMPKIGRIEGSPADRCGLKVGDRILAVNGCSITNKS 1073
 QY 863 QAVNLKNASSIEQVYVAGD-----VSVTGHH---QEPASSLSFTGLTS 907
 Db 1074 DIVNLKKEGNTVTLRIIPGDESSNATLLTNAKIATITTTTPSOQGTQETRNTPK 1133
 QY 908 TSIFQDDLCPPCK-----SITLERGPDGLFSIVGYGSPHGDULPIYKTVFAKGA 961
 Db 1134 ESQF--EKKAPQATQODEFYTELERGAKGFGFSLRG--REYMDLYVLALADGP 1188
 QY 962 EDRLKRDQIIIVNGSLGVTHAEVAAILKRTKGTVLTAV 1003
 Db 1189 RCGMKRIGDEILEINGETTKNMKHSRAILEIKNGGRVFL 1230

RESULT 12

channel associated protein of synapse 2 - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jan-2000
 C:Accession: T10811
 R:Ririe, M.; Hata, Y.; Takai, Y.
 Submitted to the EMBL Data Library, April 1996
 A:Description: Cloning of new isoforms of PSD-95/SAP90 related genes.
 A:Reference number: 217166
 A:Accession: T10811
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-852 <IR>
 A:Cross-references: EMBL:U53368; NID:g1517939; PID:g1517940
 C:Genetics:
 A:Note: PSD-95/SAP90-related gene 2 (chapsyn isoform2)
 C:Superfamily: discs-large tumor suppressor; GIGF domain homology; guanylate kinase hom
 F:198-226/Domain: GIGF domain homology <GIG>
 F:343-601/Domain: SH3 homology <SH3>
 F:663-840/Domain: guanylate kinase homology <GKI>

Query Match 8.0%; Score 407.5; DB 2; Length 852;
 Best Local Similarity 28.4%; Pred. NO. 1.7e-15;
 Matches 153; Conservative 68; Mismatches 212; Indels 111; Gaps 22;

QY 498 ENPDQAVPSAAGASG-----EKKNS-----QSLVPSGSPPESTIRNTR 541
 Db 22 DGFHDSLPRLTHEVRGPELVHSEKNSQIENHGYVQSHISPLKASPA--IIVNDT 80
 QY 542 STPAIFASDPATCPITPGCEFTIE-----ISKRTGLGLSTVGSST--LLG---AFI 592
 Db 81 LD-----TIYVNGTEIEFEFEETTLERNGSLGFSIAGCTDNHIDDDPGIFIT 130
 QY 593 EYEEGAAKQDRLMAGDOILEVNGIDLKATHDEAINVLRQTPQVRVTLTKRDAPYKE 652
 Db 131 KIIPGAAAEDERLKVNDCILNVEDVSEVSHSKAVEALKAGSTIVRLY--RRRP 189
 QY 653 EEVCDTLTLELQKPKGKGLSIVK-----RNDTGVFSDIVKGIADPDGRLIGDQ 706

Db 190 TVV-----EIKLFKGP-KGLGFSIAGVGNQIHGDNISIVTIIIDGGAOKDGRLOVGR 244
 QY 707 ILLVNGEDVRNASEAVALLKCSLCTVTLVGRKA-----GP---FHSERPQTS 756
 Db 245 LLMVNNYSLEETTHEEVAAILKNTDYYLKVKGKPTTYMTDPYGPPIITHSYSPMENN 304
 QY 757 QVSEGLSFTPLPSLSTSELES--SKKNALASEIQLRTYEMKKGPTDSLGISIVAG 816
 Db 305 LLS--GNGGLEKTSLPISPEGRSPYIPKMLV--EDETTRPE---PYSTVNLCKDP 358
 QY 817 GSPLDVPIFIAMHPTGVAQTKLRVDRIVTICGTSTEGMHTQAVNLKNASSIE 876
 Db 359 ASPRHYSPV-----ECDSF-----LLSTYPRNH 383
 QY 877 MOVVAGDVSVVYTHHQPASSLSFTGLTSTSIPODGLPCKSTYLERGPDGLFSI 936
 Db 384 LCLLPDSMTSHSQSTATROPSTVTLQRAISLE-----GEP--RKVYLHKSTGLGFI 435
 QY 937 VGGYSPHODLPIYKYTFAKGASEDRKLRDQIIIVNGSLGVTHAEVAAILKRTK 996
 Db 436 VGG---EDGE-GIFVSLIAGGPADLSELGRDQILSVNGIDLKAGASHQAAALKGAG 491
 QY 997 GTVTL 1002
 Db 492 QTVTII 497

RESULT 13

hypochemical protein C25F6.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T15617
 R:Bentley, D.
 Submitted to the EMBL Data Library, October 1995
 A:Description: The sequence of C. elegans cosmid C25F6.
 A:Reference number: 218377
 A:Accession: T15617
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1131 <BEN>
 A:Cross-references: EMBL:U39742; NID:g1049455; PID:g1049459; PIDN:AAA0434.1; CESP:C2
 A:Gene: CESP:C25F6.2
 A:Introns: 105/1; 210/3; 283/3; 316/1; 346/2; 463/3; 566/1; 722/2; 897/1; 991/2; 1032

Query Match 7.9%; Score 399.5; DB 2; Length 1131;
 Best Local Similarity 24.4%; Pred. NO. 7.4e-15;
 Matches 185; Conservative 102; Mismatches 322; Indels 149; Gaps 27;

QY 126 DIPELPEREGGEES---ELQNTAYSNMNPARRVELMRPESKLSISIVGKMGSR 181
 Db 171 EVRRVIERLEG-GPHSYNSRPAATSTSNYNLSSTPLISDLDRDGRGFSYLLNGGLGN 229
 QY 182 SMGEVMRGFTIKHYLEDSPAGKNGTLKPODRIVEAPSOSESPEKAPLCSVPPPPSAFA 241
 Db 220 GNG-----LGNGLLSPYNSST---HYLHERQRTSHGRTVETTTTTRVDPISGLE 278
 QY 242 EMGSDHTQASAKISQVDKDEDFYSWMKNIRERYGTLGLGEMIELEKSHGLSLAG 301
 Db 279 RRVVEHT-----GVDDHGRKM-----ELENYLVEKRGHGLGFSITG 315
 QY 302 NKDR---SRMSVFTVGDIPNGAAGKDRGLQIADLELLEINGOILVGRSHONASSIICAP 357
 Db 316 GMDQPTLEDGDTIYVTNIEEGALADGRMRKNDITTAANNNTNCENHVEVAVALKSSG 375
 QY 358 SKVKIIFIRKND-----AVNMAVCP-----GNAVEPLPSNSELQKKEEP 399
 Db 376 NVVSLSLRRKDEAFPLPGNGFGSTYLRSGVTSPVSAGNLQHAHISPSAPAIHPPPPP 435
 QY 400 TWTSDAAVNLSSFR---MVOHLELPKDOGLGIAIS-----EEDTSLGVIKISLT 447


```

OY 380 NAV--EPLPSNSENLOKNET-----EPTV 401
| : | | |
Db 527 LAIMERPPVPMNGRHNNTYLETYSRTSQSPDITDRPPHLSHMPADGOLDGTYPPV 586
| : | | |
OY 402 TTSDAVIDLSFRKNVOHLEPKDOGLGIAISEDTLSCVITKSLTEHGVAAITDGRKVG 461
| : | | |
Db 587 HDONVSVASSGATQALMLMTLTIVKGAKGFGFTIADSPFGQVAKQILD--IQCCPG-ICEG 643
| : | | |
OY 462 DQILAVDDIYVGYPIEKFISLKTAKM--TVKLTJHA-----ENPDS 502
| : | | |
Db 644 DLIVEINQONVONLSHTEVVDILKDCPVGSETSLIHRGFFSPMKTPKPNVDRWENQGS 703
| : | | |
OY 503 QAVPSAGAASEKKNSQSOLWPOS-----GSPEPESIRNTSRSSSTPAI 547
| : | | |
Db 704 POTSISAPV-----POSLEPPPALHRSPPDSTEAFDPKPKDPELYEKS-----AI 752
| : | | |
OY 548 FAS---DPATCPIIPGC-----ETIETISKGTGLSLVGSQD---TLGAFIHEV 594
| : | | |
Db 753 YESRQOVPRTPSRBMSDSDPKKELDVHLRMESEGFRLGDEHGPILIGA-----V 807
| : | | |
OY 595 YEEGAACAKDGRIMAGDQILEVNGIDLRKATHEBAIVLQTPQ--RVRLTLR----- 645
| : | | |
Db 808 IAMGSADRDGRLHPGDELYVDGIPVAGKTHRYVIDLMHHAARMGQVNLTVRKVLCGE 867
| : | | |
OY 646 -----DEAPYKE-----EVCDTLIELQKP 667
| : | | |
Db 868 PCPENGRSGSVSTHHSRSDYATYANSNHAAPSNNASPPGFAHSLSQTSVDYIIRKE 927
| : | | |
OY 668 GKGLSLIVGKRN--DTGVFVS-----DIVKGIADPDGRLIOGDOILLVNGEDVRNAS 719
| : | | |
Db 928 NEGFGEVLIISLNPRESGATITVPHKIGRIIDGSPADRCALKVGDRIILAVNGOSIINMP 987
| : | | |
OY 720 QEAVALKCSLCTVLEY-----GRKAGPFHSERRP--SQTQVSEGSLS 764
| : | | |
Db 988 HADIYKLIKDAGLSYTLRIIPEELNPTAPSSEKQSPMAQHSPLAQHSPLAQPSPA 1047
| : | | |
OY 765 SFTFPLSGSTSESSSKKNALASEIOGLRTVEMKKGPPTSLSIGSVGSPGLDVP 824
| : | | |
Db 1048 TENSPPVAPAPAPQPLQLOGHENSYSREVKARQDVK-----PDIRQP 1088
| : | | |
OY 825 IFIAMHPTGVAAGTQKLRVGDRIYITICGSTEGMTHQAANLKNASGSIEMQVAVAGD 884
| : | | |
Db 1089 PFTDYROP-----PLDYRQPPGCD 1107
| : | | |
OY 885 VGVVT-----GHHQEPASSLSFTGLTSTIFODDLGPPQ---CKSITLERGPDGLGFSI 936
| : | | |
Db 1108 YSQPSPLDYROHSPDTRQYPLS-----DYROPQDFDYFTVDMKKGKGFSGFSI 1155
| : | | |
OY 937 VGGYSGPHGDLPIYVKTVAFAKGAASEDRGLKRGDOIIIVNGOSLEGVTHEEVAAILKRTK 996
| : | | |
Db 1156 RGG---REYKMDLYVRLAEDGPAINGRVGDQIIEINGESTRDMTHARAIELIKSGG 1212
| : | | |
OY 997 GTVTIMV 1003
| : | | |
Db 1213 RRVRLLL 1219

```

Search completed: July 12, 2001, 14:42:06
Job time: 214 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2001, 14:43:17 ; Search time 21.71 Seconds

(without alignments) updates/sec
1585.756 Million cell

Title: US-09-502-698-2

Perfect score: 5085
Sequence: 1 MLDNVSKESEKERTINIAKGN.....EEAVAIKRTKGVITLWVLS 1005

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SWISSPROT_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	509	10.0	2485	1	PTND_HUMAN
2	405.5	8.0	852	1	DIG2_RAT
3	399	7.8	817	1	DIG2_HUMAN
4	397.5	7.8	870	1	DIG2_HUMAN
5	393	7.7	849	1	DIG3_MOUSE
6	391	7.7	849	1	DIG3_RAT
7	384	7.6	767	1	DIG4_HUMAN
8	383	7.5	724	1	DIG4_RAT
9	382	7.5	724	1	DIG4_MOUSE
10	379	7.5	960	1	DIG1_DROME
11	378.5	7.4	911	1	DIG1_RAT
12	377	7.4	904	1	DIG1_HUMAN
13	283	5.6	631	1	IL16_HUMAN
14	243	4.8	1736	1	ZOI_HUMAN
15	242	4.8	933	1	ZOI_HUMAN
16	240	4.7	898	1	ZOI_CANFA
17	237.5	4.7	1745	1	ZOI_MOUSE
18	235.5	4.6	905	1	ZOI_MOUSE
19	211	4.1	1816	1	AF6_HUMAN
20	205.5	4.0	839	1	APB1_RAT
21	182	3.6	206	1	OM25_RAT
22	177.5	3.5	571	1	APB3_MOUSE
23	177	3.5	837	1	APB1_HUMAN
24	176	3.5	145	1	OM25_HUMAN
25	173.5	3.4	569	1	APB3_RAT
26	165	3.2	982	1	ABPA_CAEEL
27	161.5	3.2	2842	1	APC_RAT
28	160	3.1	750	1	APB2_RAT
29	158	3.1	749	1	APB2_HUMAN
30	157	3.1	623	1	ADH_DROME
31	153.5	3.0	2845	1	MAPB_MOUSE
32	153	3.0	2468	1	MAPB_HUMAN
33	151	3.0	680	1	APB2_MOUSE

ALIGNMENTS

RESULT 1
PTND_HUMAN STANDARD: PRT: 2485 AA.
AC Q12923: Q15263: Q16826: Q15264: Q15265: Q15159;
AD 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROTEIN TYROSINE PHOSPHATASE, NON-RECEPTOR TYPE 13 (EC 3.1.3.48)
DE (PROTEIN-TYROSINE PHOSPHATASE 1E) (PTP-RT) (PTP-BAS) (PROTEIN-TYROSINE PHOSPHATASE PTP1) (FAS-ASSOCIATED PROTEIN-TYROSINE PHOSPHATASE 1) (FAP-1).
DE PTPN13 OR PTP1E OR PTP1 OR PNP1.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_Taxid-9606;
OX [1]
RN TISSUE-Breast carcinoma;
RC MEDLINE-94350988; PubMed-8071359;
RA Banville D., Ahmed S., Stocco R., Shen S.-H.;
RT "A novel protein-tyrosine phosphatase with homology to both the cytoskeletal proteins of the band 4.1 family and junction-associated guanylate kinases";
RT J. Biol. Chem. 269:22320-22327(1994).
RT [2]
RN TISSUE-Breast carcinoma;
RC MEDLINE-94116679; PubMed-8287977;
RA Maekawa K., Imagawa N., Nagamatsu M., Harada S.;
RT "Molecular cloning of a novel protein-tyrosine phosphatase containing a membrane-binding domain and GIGF repeats";
RT FEBS Lett. 337:200-206(1994).
RN [3]
RN TISSUE-Breast carcinoma;
RC MEDLINE-95014139; PubMed-7929060;
RA Sares J., Claesson-Welsh L., Heldin C.-H., Gonen L.J.;
RT "Cloning and characterization of PTP1, a protein tyrosine phosphatase with similarities to cytoskeletal-associated proteins";
RT J. Biol. Chem. 269:24082-24089(1994).
RN [4]
RN TISSUE-Pancreas;
RC MEDLINE-95014139; PubMed-7929060;
RA Wang H.Y.;
RT "SEQUENCE OF 1216-2490 FROM N.A.
RN Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
RN [5]
RN STRUCTURE BY NMR OF 1361-1456.
RX MEDLINE-20170882; PubMed-10704206;
RA Kozlov G., Gehring K., Ekiel I.;
RT "Solution structure of the PDZ domain from human phosphatase hPTP1E and its interactions with C-terminal peptides from the Fas receptor";
RT Biochemistry 39:2572-2580(2000).
CC -I- FUNCTION: BINDS TO A NEGATIVE REGULATORY DOMAIN IN FAS THAT INHIBITS FAS-INDUCED APOPTOSIS.

OY 845 GDRIVTIGTSTEGMTHQAVNLKKNASSIEMQVAGGVSVYTHGHOE-PASSLSFT 903
 DB 1831 GDRLKVDVTDVNTMTHTDANLRAASTVNL-----VGRVLELPRIML--- 1877
 OY 904 GLTSTISODDLGPPQCKSTLERGPDGLGFSIVGSGPHGDLIVYKTFPAKASSED 963
 DB 1878 -----PHLLPDITLTCKNEELGFSLGGHSLV--QVYVSDIMPRSVAAIE 1922
 OY 964 GLKRGDOIIAVNGOSLSECVTHEEA-----VALKRTKGYVTLM 1002
 DB 1923 GNLQILDVHYVNGVSTQGMLEEVNRALDMSLPVLVTKATRNDDLPV 1970

RESULT 2
 DLG2_RAT 2
 ID DLG2_RAT STANDARD: PRT: 852 AA.
 AC Q63622; Q62939; P70548;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE CHANNEL ASSOCIATED PROTEIN OF SYNAPSE-110 (CHASYN-110) (SYNAPTIC DENSITY PROTEIN PSD-93) (DISCS, LARGE HOMOLOG 2).
 GN DLG2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96310881; PubMed=8755482;
 RA Kim E., Cho K.O., Rothschild A., Sheng M.;
 RT "Heteromultimerization and NMDA receptor-clustering activity of Chapsyn-110, a member of the PSD-95 family of proteins.";
 RL Neuron 17:103-113(1996).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96193770; PubMed=8625413;
 RA Brennan J.E., Chao D.S., Gee S.H., McGee A.W., Craven S.E., Santillano D.R., Wu Z., Huang F., Xia H., Peters M.F.,
 RA Froehner S.C., Bredt D.S.;
 RT "Interaction of nitric oxide synthase with the postsynaptic density protein PSD-95 and alpha1-syntrophin mediated by PDZ domains.";
 RL Cell 84:757-767(1996).
 RN (3)
 RP SEQUENCE FROM N.A.
 RA Irie M., Hata Y., Takai Y.;
 RL Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR SUBUNITS AS WELL AS POTASSIUM CHANNELS.
 CC -1- SIMILARITY: CONTAINS 3 PDZ/DRH DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@sib-sib.ch).
 CC -----
 CC EMBL: U49049; AAB53243.1; -
 DR EMBL: U50717; AAC52643.1; -
 DR EMBL: U53368; AAB48562.1; -
 DR HSSP: Q12959; 1PDR.
 DR InterPro: IPR000619; -
 DR InterPro: IPR001452; -
 DR InterPro: IPR001478; -
 DR Pfam: PF00625; Guanylate_kin; 1.
 DR Pfam: PF00595; PDZ; 3.
 DR Pfam: PF00018; SH3; 1.
 DR PROSITE: PS00856; GUANYLATE_KINASE_1; 1.

DR PROSITE: PS0052; GUANYLATE_KINASE_2; 1.
 DR PROSITE: PS0106; PDZ; 3.
 DR PROSITE: PS0002; SH3; 1.
 KM SH3 domain; Repeat.
 FT DOMAIN 98 184
 FT DOMAIN 193 279
 FT DOMAIN 421 501
 FT DOMAIN 536 606
 FT DOMAIN 662 852
 FT CONFLICT 181 182
 FT CONFLICT 228 228
 FT CONFLICT 326 326
 FT CONFLICT 339 339
 FT CONFLICT 450 454
 FT CONFLICT 464 465
 FT CONFLICT 474 474
 FT CONFLICT 476 476
 FT CONFLICT 478 478
 FT CONFLICT 484 486
 FT CONFLICT 506 506
 FT CONFLICT 569 569
 FT CONFLICT 586 586
 FT CONFLICT 626 641
 FT CONFLICT 639 639
 FT CONFLICT 726 726
 FT CONFLICT 733 733
 FT CONFLICT 749 749
 FT CONFLICT 756 756
 FT CONFLICT 791 792
 FT CONFLICT 794 794
 SO SEQUENCE 852 AA; 94934 MM; FBD414A8B9CF5B09 CnC64;
 T -> M (IN REF. 2).
 KR -> NG (IN REF. 2).
 L -> H (IN REF. 2).
 D -> H (IN REF. 2).
 R -> P (IN REF. 2).
 A -> D (IN REF. 2).
 AAA -> LP (IN REF. 2).
 GD -> RK (IN REF. 2).
 MISSING (IN REF. 2).
 MISSING (IN REF. 2).
 MISSING (IN REF. 2).
 A -> S (IN REF. 2).
 H -> N (IN REF. 2).
 L -> Q (IN REF. 2).
 MISSING (IN REF. 2).
 K -> A (IN REF. 3).
 F -> L (IN REF. 1).
 N -> Y (IN REF. 2).
 E -> V (IN REF. 1).
 L -> H (IN REF. 2).
 KR -> NG (IN REF. 2).
 T -> M (IN REF. 2).

Query Match 8.0%; Score 405.5; DB 1; Length 852;
 Best Local Similarity 28.4%; Pred. No. 9.9e-15;
 Matches 155; Conservative 67; Mismatches 213; Indels 111; Gaps 22;

OY 498 ENPDQAVPSAAGASG-----EKNSS-----QSLMVGSGSPPEESIRNRS 541
 DB 22 DGPHDHSLRLTHEVGRPELVHSEKNSQIENHGVYSHISPLKASPAE-IIVNDR 80
 OY 542 SSTPAIFADPATCPPIPCETTIE-----ISKRTGLGSLVSGSDT-LLG---AFITH 592
 DB 81 LD-----TIPVNGTEIEFEETLERGSGSLGFIAGTDNPHIGDDPCIFIT 130
 OY 593 EYERGAACKDRLWAGDOILEVNGIDLRKATHEAIVNLRQTPQVRVRLTYRDEAPYKE 652
 DB 131 KIIPGAAAEEDRLRNDCILRNVEDVSEVSHKAVELAKGASIVRLYV-RRRRPIL 189
 OY 653 EBYCDTLTIELOKKPKGKGLSIVK-----RNDGVFSDIVKGIADPPGRLIQDQ 706
 DB 190 TVV-----EIKLRKGP-KGIGFSIAGVGNGHIGPNSIYVTKIIDGAAOKGKRLQVGR 244
 OY 707 ILLVNGEDVRNMSOEVAVALKCSLGTVLVEYGRKA-----GP---FHSERPSQTS 756
 DB 245 LLMVNNYSLEEVTHEEVAAILKNTSDVYLKKGKPTTIYMPYGPDPDTHSISPMENH 304
 OY 757 QVSEGLSFTPEPLSSSTSESLSSSKNALASEIOLRTVEEMKGPDSLGISAGVY 816
 DB 305 LLS-GNNGTLEKXTSLPISPRYSPIPKHMLV-EDDYRPRPE---PYSTVNNLCQDP 358
 OY 817 GSPPLGDVPIFFIMMHPITGVAAOTOKLRVGDRLVITTCGSTBECHTHQAVNLKKNASGSTE 876
 DB 359 ASPRHYSFV-----EDCKSF-----LSTPYPHYH 383
 OY 877 MQVVAAGDVSVYTHGHOEPASSLSFTGLTSTISFODDLGPPQCKSTLERGPDGLGFSI 936
 DB 384 LGLLPDSWNTSHSQSTATROPSVTLQRAISLE-----GEP--RKVYLHKSGTGLGFRYI 435
 OY 937 VCGYSGPHGDLPIYKYTFPAKGAASEDRGLKRGDOIIAVNGOSLSECVTHEEVAAILKRTK 996
 DB 436 VGG---EDGE-GIFVSFLIAGGPADLSGELQRGDQILSVNGIDLGASHEQAAALAKGAG 491

ID	NAME	STANDARD	PRT	817 AA
DLG3_HUMAN				
092796	OSU118			
01-NOV-1997	(Rel. 35, Created)			
01-NOV-1997	(Rel. 35, Last sequence update)			
01-OCT-2000	(Rel. 40, Last annotation update)			
PRESYNAPTIC PROTEIN SAP102 (SYNASE-ASSOCIATED PROTEIN 102)				
(NEUROENDOCRINE-DLG) (NE-DLG) (DISCS, LARGE HOMOLOG 3).				
DLG3 OR KIAA1232.				
Homo sapiens (Human).				
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
NCBI_TaxID=9606;				
[1]				
SEQUENCE FROM N.A.				
TISSUE=Fetal brain;				
MEDLINE=97332623; PubMed=9188857;				
Makino K., Kuwahara H., Masuko N., Nishiyama Y., Morisaki T.,				
Sasaki J., Nakao M., Kiwano A., Nakata M., Ushio Y., Saya H.;				
"Cloning and characterization of NE-dlg, a novel human homolog of the				
Drosophila discs large (dlg) tumor suppressor protein interacts with				
the APC protein.";				
Oncogene 14:2425-2433(1997).				
[2]				
SEQUENCE OF 330-817 FROM N.A.				
TISSUE=Brain;				
MEDLINE=20039619; PubMed=10574462;				
Nagase T., Ishikawa K.-I., Kikuno R., Hirotsawa M., Nomura N.,				
Ohara O.;				
"Prediction of the coding sequences of unidentified human genes. XV.				
The complete sequences of 100 new cDNA clones from brain which code				
for large proteins in vitro."				
DNA Res. 6:337-345(1999).				
-1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF THE NMDA RECEPTOR				
SUBUNIT NR2B (BY SIMILARITY).				
-1- SIMILARITY: CONTAINS 3 PDZ/DRH DOMAINS.				
-1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.				
-1- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.				
-1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.				

This SWISS-PROT entry is copyright. It is produced through a collaboration				
between the Swiss Institute of Bioinformatics and the EMBL outstation -				
the European Bioinformatics Institute. There are no restrictions on its				
use by non-profit institutions as long as its content is in no way				
modified and this statement is not removed. Usage by and for commercial				
entities requires a license agreement (See http://www.isb-sib.ch/announce/				
or send an email to license@sib-sib.ch).				

EMBL; U49089; AAB61453.1; -.				
EMBL; AB033058; BAA6546.1; -.				
HSSP; Q12959; 1PDR.				
MIM; 300189; -.				
InterPro; IPR000619; -.				
InterPro; IPR001452; -.				
InterPro; IPR001478; -.				
Pfam; PF006623; Guanylate_kin; 1.3				
Pfam; PF00595; PDZ; 3.				
Pfam; PF00018; SH3; 1.				
PROSITE; PS00856; GUANYLATE_KINASE_1; 1.				
PROSITE; PS50052; GUANYLATE_KINASE_2; 1.				
PROSITE; PS50106; PDZ; 3.				
PROSITE; PS50002; SH3; 1.				
SH3 domain: Repeat.				
DOMAIN 130 217 PDZ 1.				
FT DOMAIN 226 311 PDZ 2.				
FT DOMAIN 379 465 PDZ 3.				
FT DOMAIN 503 568 SH3.				

[illegible]

ID	DLG4_HUMAN	STANDARD;	PRF;	767 AA.
AC	P78352; Q92941; Q90RK8;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	03-MAY-2000 (Rel. 39, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	PRESYNAPTIC DENSITY PROTEIN 95 (PSD-95) (DISCS, LARGE HOMOLOG 4).			
GN	DLG4 OR PSD95.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RX	MEDLINE=97432822; PubMed=9286702;			
RA	Stethakis D.G., Hoover K.B., You Z., Bryant P.J.;			
RT	"Human postsynaptic density-95 (PSD95): location of the gene (DLG4)			
RL	and possible function in nonneural as well as in neural tissues.";			
RL	Genomics 44:71-82(1997).			
RN	[2]			
RP	REVISIONS.			
RC	TISSUE=Mammary gland;			
RA	Stethakis D.G., Hoover K.H., You Z., Bryant P.J.;			
RL	Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Stethakis D.G., Udar N., Sandgren O., Andreasson S., Bryant P.J.,			
RT	Small K., Forsman-Semb K.;			
RL	"Genomic organization of human DLG4, the gene encoding post-synaptic			
RT	density 95 (PSD95).";			
RL	J. Neurochem. 73:0-0(1999).			
RN	[4]			
RP	SEQUENCE OF 81-401 FROM N.A.			
RC	TISSUE=Brain;			
RA	Brennan J.E., Bredt D.S., Parkinson J.F., Manzana W.P., McClary J.A.;			
RL	Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.			
CC	-1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR			
CC	SUBUNITS. MAY BE INVOLVED IN SYNAPTOGENESIS.			
CC	-1- SUBCELLULAR LOCATION: CONCENTRATED AT SYNAPTIC JUNCTIONS PRIMARILY			
CC	ON THE PRESYNAPTIC SIDE (WAS ORIGINALLY THOUGHT TO BE			
CC	POSTSYNAPTIC).			
CC	-1- TISSUE SPECIFICITY: PRESYNAPTIC DENSITY FRACTION OF BRAIN.			
CC	-1- SIMILARITY: CONTAINS 3 PDZ/DRR DOMAINS.			
CC	-1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.			
CC	-1- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.			
CC	-1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on ways			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL; U83192; AAC52113.1; -			
DR	EMBL; AF156495; AAD56173.1; -			
DR	EMBL; U68138; AAB07736.1; -			
DR	HSSP; Q12959; 1PDR.			
DR	MIM; 602887; -			
DR	InterPro; IPR000619; -			
DR	InterPro; IPR001452; -			
DR	InterPro; IPR001478; -			
DR	PIfam; PF00595; PDZ; 3.			
DR	PIfam; PF00018; SH3; 1.			
DR	PROSITE; PS00856; GUANYLATE_KINASE_1; 1.			
DR	PROSITE; PS50052; GUANYLATE_KINASE_2; 1.			
DR	PROSITE; PS50106; PDZ; 3.			
DR	PROSITE; PS50002; SH3; 1.			
KW	SH3 domain; Repeat.			

FT	DOMAIN	108	194	PD2 1.
FT	DOMAIN	203	289	PD2 2.
FT	DOMAIN	356	436	PD2 3.
FT	DOMAIN	471	541	SH3.
FT	DOMAIN	577	767	GUNNATE KINASE.
FT	CONFLICT	46	46	E -> V (IN REF. 3).
FT	CONFLICT	81	83	VIV -> EFR (IN REF. 4).
FT	CONFLICT	399	401	GDO -> AGI (IN REF. 4).
SO	SEQUENCE	767 AA;	85429 MW;	BE1019159E65B2D8 CRC4;

Query Match	7.68;	Score 384;	DB 1;	Length 767;
Best Local Similarity	25.88;	Pred. NO. 1.2e-13;		
Matches 131;	Conservative 76;	Mismatches 153;	Indels 148;	Gaps

QY	518	NSOSLWYPOQSGP----	EPEISINTSKSSSTPAIFASDPATCP----	IIFCETITIE--	566
Db <td>50</td> <td>SESKYRYRQDDPTPLRPSAHLNQA--NSPPIVNTDITLAPQELQVNGTSEMEYEE</td> <td></td> <td></td> <td>108</td>	50	SESKYRYRQDDPTPLRPSAHLNQA--NSPPIVNTDITLAPQELQVNGTSEMEYEE			108
QY <td>567</td> <td>--ISKRTGCLGLSTVSGSDT--LLG----<td>AFITHEVYEGCAACKGRMLAGDQILEVNGIDL</td><td>620</td><td>620</td></td>	567	--ISKRTGCLGLSTVSGSDT--LLG---- <td>AFITHEVYEGCAACKGRMLAGDQILEVNGIDL</td> <td>620</td> <td>620</td>	AFITHEVYEGCAACKGRMLAGDQILEVNGIDL	620	620
Db <td>109</td> <td>ITLEGNNGSLGFGSIAGCGDNDHIDDDISFTFKIIPGGAADRLRVNDSITLFVNEVDV</td> <td></td> <td></td> <td>168</td>	109	ITLEGNNGSLGFGSIAGCGDNDHIDDDISFTFKIIPGGAADRLRVNDSITLFVNEVDV			168
QY <td>621</td> <td>RKATDEAIIVLRQTPQGVRLTLKRDAPYKEEVCOTLTITELQKRPKGKGLSTYVK--</td> <td></td> <td></td> <td>678</td>	621	RKATDEAIIVLRQTPQGVRLTLKRDAPYKEEVCOTLTITELQKRPKGKGLSTYVK--			678
Db <td>169</td> <td>REVTHSAVVELKENGSTIVRLTYMKRRPRAEK-----VMEIKLIKGP-KGLGFSIAGCVG</td> <td></td> <td></td> <td>222</td>	169	REVTHSAVVELKENGSTIVRLTYMKRRPRAEK-----VMEIKLIKGP-KGLGFSIAGCVG			222
QY <td>679</td> <td>----RNDTGVVSDIVKGGIADPPGRRLIQDQIILLVNGEDVRNASEVAALAKCSLGTV</td> <td></td> <td></td> <td>734</td>	679	----RNDTGVVSDIVKGGIADPPGRRLIQDQIILLVNGEDVRNASEVAALAKCSLGTV			734
Db <td>223</td> <td>NQHTPGDMSIYVTIKIEGAAHKMGKRLQIGKILAVNSVGLIEDVMHEDAVAAALKNTYDV</td> <td></td> <td></td> <td>282</td>	223	NQHTPGDMSIYVTIKIEGAAHKMGKRLQIGKILAVNSVGLIEDVMHEDAVAAALKNTYDV			282
QY <td>735</td> <td>TLENGRIKAGFPFHSERPSOTSQVSEGLSFTPLPSGSSSTSESLESSKKNALASEIQG</td> <td></td> <td></td> <td>794</td>	735	TLENGRIKAGFPFHSERPSOTSQVSEGLSFTPLPSGSSSTSESLESSKKNALASEIQG			794
Db <td>283</td> <td>YLKVA-----KPS-NAYLS-----SYAPPDITTSYSHLDNEISHSS-----</td> <td></td> <td></td> <td>319</td>	283	YLKVA-----KPS-NAYLS-----SYAPPDITTSYSHLDNEISHSS-----			319
QY <td>795</td> <td>LRTYMKKGPPLDSLGISLAGVGSPLGDVPEFIIMMHPGTGAQTQRLRGDRIVTICGT</td> <td></td> <td></td> <td>854</td>	795	LRTYMKKGPPLDSLGISLAGVGSPLGDVPEFIIMMHPGTGAQTQRLRGDRIVTICGT			854
Db <td>320</td> <td>-----YLGDTPIPT-----</td> <td></td> <td></td> <td>327</td>	320	-----YLGDTPIPT-----			327
QY <td>855</td> <td>STEGTHTQAVNLTKNAGSIEQMVGAGDVSVTGHHQEPASSLSFTGLTSTIPQDD</td> <td></td> <td></td> <td>914</td>	855	STEGTHTQAVNLTKNAGSIEQMVGAGDVSVTGHHQEPASSLSFTGLTSTIPQDD			914
Db <td>328</td> <td>---AMTPT-----SPRYSVPAKULGEED</td> <td></td> <td></td> <td>349</td>	328	---AMTPT-----SPRYSVPAKULGEED			349
QY <td>915</td> <td>LGPPOCKSITLERPDDLGFSIVSGYSGPHGLDPIYKTYFAKCAASEDGRLKRGDQIIA</td> <td></td> <td></td> <td>974</td>	915	LGPPOCKSITLERPDDLGFSIVSGYSGPHGLDPIYKTYFAKCAASEDGRLKRGDQIIA			974
Db <td>350</td> <td>I-PPRPRIIVHRSGTSGFNFVGG---EDGE-GIFISFIAGSPADLSGELRGKQDILS</td> <td></td> <td></td> <td>404</td>	350	I-PPRPRIIVHRSGTSGFNFVGG---EDGE-GIFISFIAGSPADLSGELRGKQDILS			404
QY <td>975</td> <td>VNGOSLEGVTHEAVAILLKRTKRGVTLM</td> <td></td> <td></td> <td>1002</td>	975	VNGOSLEGVTHEAVAILLKRTKRGVTLM			1002
Db <td>405</td> <td>VNGVDLRNASHHQAAIALKNAGQVTTII</td> <td></td> <td></td> <td>432</td>	405	VNGVDLRNASHHQAAIALKNAGQVTTII			432

RESULT	8
DIG4_RAT	STANDARD; PRT: 724 AA.
AC	P31016; P97631; 26, Created
DT	01-JUL-1993 (Rel. 26, Last sequence update)
DT	01-JUL-1993 (Rel. 26, Last sequence update)
DT	01-OCT-2000 (Rel. 40, Last annotation update)
DE	PRESYNAPTIC DENSITY PROTEIN 95 (PSD-95) (PRESYNAPTIC PROTEIN SAP90)
DE	(STNABSE-ASSOCIATED PROTEIN 90) (DISCS, LARGE HOMOLOG 4).
GN	DIG4 OR PSD95.
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus
OX	NCBI_TaxID=10116;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RX	MEDLINE=93040233; PubMed=1419001;
RA	Cho K.-O., Hunt C.A., Kennedy M.B.;
RT	"The rat brain postsynaptic density fraction contains a homolog of the Drosophila discs-large tumor suppressor protein.";

```

RL  Neuron 9:929..942(1992).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN-SPRAGUE-DAWLEY; TISSUE=Brain;
RX  MEDLINE=93186749; PubMed=7680343;
RA  Kistner U., Wenzel B.M., Voh R.W., Cases-Ingloff C., Garner A.M.,
RT  Appelbauer U., Voss B., Gundelfinger E.D., Garner C.C.;
RT  "SAP90, a rat presynaptic protein related to the product of the
RT  Drosophila tumor suppressor gene dlg-A.";
RL  J. Biol. Chem. 268:4580-4583(1993).
RN  [3]
RP  SEQUENCE OF 566-625 FROM N.A.
RC  STRAIN=Wistar Kyoto; TISSUE=vascular smooth muscle;
RA  Adams L.D., Werny I., Schwartz S.M.;
RL  Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN  [4]
RP  X-RAY CRYSTALLOGRAPHY (1.82 ANGSTROMS) OF 302-402.
RX  MEDLINE=96270509; PubMed=8674113;
RA  Doyle D.A., Lee A., Lewis J., Kim E., Sheng M., Mackinnon R.;
RT  "Crystal structures of a complexed and peptide-free membrane protein-
RT  binding domain: molecular basis of peptide recognition by PDZ".
RL  Cell 85:1067-1076(1996)
CC  -1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR
CC  SUBUNITS. MAY BE INVOLVED IN SYNAPTOGENESIS.
CC  -1- SUBCELLULAR LOCATION: CONCENTRATED AT SYNAPTIC JUNCTIONS PRIMARILY
CC  ON THE PRESYNAPTIC SIDE (WAS ORIGINALLY THOUGHT TO BE
CC  POSTSYNAPTIC).
CC  -1- TISSUE SPECIFICITY: PRESYNAPTIC DENSITY FRACTION OF BRAIN.
CC  -1- SIMILARITY: CONTAINS 3 PDZ/DRH DOMAINS.
CC  -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC  -1- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
CC  -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC  -----
CC  CC This SWISS-PROT entry is copyrigh. It is produced through a collaboration
CC  CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  CC the European Bioinformatics Institute. There are no restrictions on its
CC  CC use by non-profit institutions as long as its content is in no way
CC  CC modified and this statement is not removed. Usage by and for commercial
CC  CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC  CC or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL, M96853; AAA41971.1; -
DR  EMBL, X66474; CAA47103.1; -
DR  EMBL, U77090; AAB38270.1; -
DR  PIR, S26407; S26407.
DR  PIR, JH0800; JH0800.
DR  PDB, 1BE9; 2I-OCT-98.
DR  PDB, 1BFE; 2I-OCT-98.
DR  InterPro: IPR000619; -
DR  InterPro: IPR001452; -
DR  InterPro: IPR001478; -
DR  Pfam: PF00625; Guanylate_kin; 1.
DR  Pfam: PF00595; PDZ; 3.
DR  Pfam: PF00018; SH3; 1.
DR  PROSITE: PS00856; GUANYLATE_KINASE_1; 1.
DR  PROSITE: PS50052; GUANYLATE_KINASE_2; 1.
DR  PROSITE: PS50106; PDZ; 3.
DR  PROSITE: PS50002; SH3; 1.
KW  SH3 domain; Repeat; 3D-structure.
FT  DOMAIN 65..151 PDZ 1.
FT  DOMAIN 160..246 PDZ 2.
FT  DOMAIN 313..393 PDZ 3.
FT  DOMAIN 428..498 SH3.
FT  DOMAIN 534..724 GUANYLATE KINASE.
FT  CONFLICT 61..61 M->L (IN REF. 2).
FT  CONFLICT 78..78 S->T (IN REF. 2).
FT  CONFLICT 177..182 G->G (IN REF. 2).
FT  CONFLICT 200..200 AVGNQH->ALGTSI (IN REF. 2).
FT  CONFLICT 254..254 S->T (IN REF. 2).
FT  CONFLICT 539..555 ILGPTKQANDDLSEF->ISLDPKTPTVMTIFSSS
FT  (IN REF. 2).
FT  CONFLICT 623..625 GKH->RDO (IN REF. 3).
FT  SEQUENCE 724 AA: 80465 MW: 792204EBE0F9AD85 CCK64;

```

```

Query Match Similarity 7.5%: Score 383; DB 1; Length 724;
Best Local Similarity 26.2%: Pred No. 1,3e-13;
Matches 128; Conservative 74; Mismatches 143; Indels 144; Gaps 17.

OY 533 PESIRNTSSSTPAIFASDPATCP-----IIPCETTIE-----ISKRTGLSLSTVGSD 583
DB 141 L P N O A : N S P P V I N T D T L E A P G E I L Q V N T E G E M E Y E I T L E R G N G L F S F I N G T D 84
OY 584 T L L G --- A F I T H E Y E E B A A C K D R G L Y A M G D O I L E V N G I D L K R A T H D E L I N V L R O T P O V 639
DB 85 N P H I G D P S I F T K K I I P G G A A A O D G L R V N S I L F E V N E Y D V R E T H S A V E A L K E N G S I V 144
OY 640 R L T L Y R D E A P Y K E E E V C D L T T E L Q K K P G K G L S I V K C ----- R N D G V F V S D I V K G 693
DB 145 R L Y V M R K R P A K ----- V M E I K L I K G P - K G L G F S I A G V G N O H I P G D N S I Y T K I I E G G 198
OY 694 I A D P G R L I Q G D O I L L V N G E D V R N M S O E A V A A L L C S I G T V T L E V G R I K A G P H S E R R P S 753
DB 199 A A H K G R L D I G K I I A V N S V G L E D V H E D A V A A L K N T Y D V Y L K V A ----- R P S 247
OY 754 Q R S O V S E G S L S F T P L S S S T S E S L S S K K N A L S E I Q G L R Y E M K K P T D S I G I S T A 813
DB 248 - N A V I S D ----- S Y A P D I T T S Y Q H I D N E I S H S ----- 276
OY 814 G G V G S P L G D V P I F I A M H P T G V A A O T O K R L V R D R I Y T I C G T S T E G M T H T O A V N L K M A S G 873
DB 277 ----- Y L G D Y P T ----- A M T P T ----- 289
OY 874 S I E M Q V A G D V S Y T T G H H G P A S S L S T G T S I S F I O D D L G P P O C K S I T L E R P D I G 933
DB 290 ----- S P R A I S P A K D L G E E D I - P E P P R I V I H R S T G S 324
OY 934 F S I V G Y G S P H G D L P I Y V T F A K G A S E D G R L K R G D I I A V N G S L E G V T H E E A V A I L K 993
DB 325 F N I V G S --- E D G E - G I F I S F I L A G P A D L S G L R G D O L S I N G V D L R N A S H E Q A A I A L K 380
OY 994 R T K G V T T L M 1002
DB 381 N A G O T V T T I 389

RESULT 9
DLG4_MOUSE
AC 062108; STANDARD; PRT; 724 AA.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DI 01-OCT-2000 (Rel. 40, Last annotation update)
DE PRESYNAPTIC DENSITY PROTEIN 95 (PSD-95) (PRESYNAPTIC PROTEIN SAP90)
DE (SYNAPSE-ASSOCIATED PROTEIN 90) (DISCS, LARGE HOMOLOG 4).
GN DLG4 OR DLG4 OR PSD95.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBA/2; TISSUE=Brain;
RL Submitted (MAY-1995) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR
CC SUBUNTS. MAY BE INVOLVED IN SYNAPTOGENESIS (BY SIMILARTY).
CC -1- SUBCELLULAR LOCATION: CONCENTRATED AT SYNAPTIC JUNCTIONS PRIMARILY
CC ON THE PRESYNAPTIC SIDE (BY SIMILARTY).
CC -1- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sdb.ch/announce/> or send an email to license@isb-sdb.ch).

DR EMBL: D50621; BAA09297.1; -
DR HSSP: Q12959; 1PDR.
DR MGP: MGI:1277959; D1gh4.
DR InterPro: IPR000619; -
DR InterPro: IPR001478; -
DR Pfam: PF00625; Guanylate_kin; 1.
DR Pfam: PF00595; PDZ; 3.
DR Pfam: PF00018; SH3; 1.
DR PROSITE: PS00856; GUANYLATE_KINASE_1; 1.
DR PROSITE: PS50052; GUANYLATE_KINASE_2; 1.
DR PROSITE: PS50106; PDZ; 3.
DR PROSITE: PS50002; SH3; 1.
KW SH3 domain; Repeat.
FT DOMAIN 65 151 PDZ 1.
FT DOMAIN 160 246 PDZ 2.
FT DOMAIN 313 393 PDZ 3.
FT DOMAIN 428 498 SH3.
FT DOMAIN 534 724 GUANYLATE KINASE.
SQ SEQUENCE 724 AA; 80472 MW; 7EFEC99E1FF90BA CRC64;

Query Match 7.5%; Score 382; DB 1; Length 724;
Best Local Similarity 26.2%; Pred. No. 1.5e-13;
Matches 128; Conservative 74; Mismatches 143; Indels 144; Gaps 17;

OY 533 PESINTSSSTPAFAFASPAFCP-----IIPGCEITIE-----ISKRGIGLGLSTVGSD 583
DB 26 PAHLNQA--NSPPVIVNTDLEAPGEYLOVNGTEGMEYEITLLENGNGSLFGSTAGTD 84
OY 584 T-LIG---AFITHEVEEGAACKDGRMLAGDQILEVNGIDLRKATHEAIVNLRQTPQAV 639
DB 85 NHIDDDPSIFITKIIIPGGAADGRLRYNDISILFVNEVDVETHSNAVELKENGSLV 144
OY 640 RLTYRDEAPYKEEYVCDLTLELQKKPKGLGLSTVGK-----RNDTGVFVSDIVKG 693
DB 145 RLYVMRRKPAEK-----IIEIKLIKGP-KGLGFSIAGGVGNQHLPGDINSIYVTKIIEGG 198
OY 694 IADPGRLQGOIILLVNEEDVRNMQEVAALLKCSLGTVLENGRIAGFHSRRS 753
DB 199 AAHKGRLQIGKILAVNSVGLVDHEDAVALKNTYDVYIKVA-----KPS 247
OY 754 QTSQVSEGLSFTPLSGSTSESSSKKNALASEIQLRTVEMKGPIDSLGISIA 813
DB 248 -NAVLSD---SYAPPDITTSYQHLDNEISHS----- 276
OY 814 GGVSPFLGDVPFIAMMHPVAAOTOKLRVDRIVTICGTSTEGMTHQAVNLKNASG 873
DB 277 -----YLGTDYPT-----AMTPT----- 289
OY 874 SIEMQVAGGVSVYTGHHQPAASSLSLTGLSTSIIPDDDLGPPCKRITLERGPDGIG 933
DB 290 -----SPRRSPVAKDLGLEDI-PREPRIVIHKRGSTG 324
OY 934 FSIVGGYSGPHGDLPIYKVFKAASDGRRLKRGDOIIVANGSLSEVTHEEVAAILK 993
DB 325 FNIIVG---EDE-GIFISFILAGCPADISGELRKDDQLTSVNGVDLRNASHQAAIALK 380
OY 994 RTKGVTTLM 1002
DB 381 NAGGVYTI 389

RESULT 10
DLG1_DROME STANDARD; PRT; 960 AA.
ID DLG1_DROME
AC P31007;

DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE LETHAL(1)DISCS LARGE-1 TUMOR SUPPRESSOR PROTEIN.
GN DLG1.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP MEDLINE=91330294; PubMed=1651169;
RX Woods D.F., Bryant P.J.;
RT "The discs-large tumor suppressor gene of Drosophila encodes a
RT guanylate kinase homolog localized at septate junctions.";
RL Cell 66:451-464(1991).
CC -1- FUNCTION: LOSS OF THIS PROTEIN CAUSES LARGE IMAGINAL DISKS BY
CC ALLOWING NEOPLASTIC OVERGROWTH. THE PRESENCE OF A GUANYLATE KINASE
CC DOMAIN SUGGESTS INVOLVEMENT IN ADEQUATE CELLULAR ADHESION AS WELL
CC AS SIGNAL TRANSDUCTION TO CONTROL CELLULAR PROLIFERATION. THE DLG
CC PROTEIN PLAYS A CRITICAL ROLE AT SEPTATE JUNCTIONS IN CELLULAR
CC GROWTH CONTROL DURING LARVAL DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: CYTOSKELETON-ASSOCIATED. LOCATED AT THE
CC CYTOPLASMIC FACE OF THE MEMBRANE IN THE CELLULAR BLASTODERM AND
CC BECOMES ASSOCIATED WITH SEPTATE JUNCTIONS WHICH BEGIN TO FORM
CC BETWEEN EPITHELIAL CELLS AT THE TIME OF DORSAL CLOSURE. IN ADULT
CC FLIES, LOCATED AT THE APICAL-LATERAL MEMBRANE BOUNDARY OF
CC EPITHELIAL CELLS.
CC -1- ALTERNATIVE PRODUCTS: LEADS TO DIFFERENT FORMS OF THE PROTEIN;
CC THE SEQUENCE SHOWN REFERS TO THE DLG-A PROTEIN, THAT IS MOST
CC ABUNDANTLY EXPRESSED.
CC -1- TISSUE SPECIFICITY: BROAD SPECIFICITY. ABUNDANTLY EXPRESSED IN
CC EPITHELIAL TISSUES. DEVELOPMENTAL STAGE DEPENDENT EXPRESSION.
CC -1- SIMILARITY: CONTAINS 3 PDZ/DRH DOMAINS.
CC -1- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
CC or send an email to license@isb-sdb.ch).

DR EMBL: M73529; AAA28468.1; -
DR PIR: A39651; A39651.
DR HSSP: Q12959; 1PDR.
DR Flybase: FBgn0001624; dlgl.
DR InterPro: IPR000619; -
DR InterPro: IPR001478; -
DR Pfam: PF00625; Guanylate_kin; 1.
DR Pfam: PF00595; PDZ; 3.
DR Pfam: PF00018; SH3; 1.
DR PROSITE: PS00856; GUANYLATE_KINASE_1; 1.
DR PROSITE: PS50052; GUANYLATE_KINASE_2; 1.
DR PROSITE: PS50106; PDZ; 3.
DR PROSITE: PS50002; SH3; 1.
KW Transducer; SH3 domain; Alternative splicing; Repeat.
FT DOMAIN 40 126 PDZ 1.
FT DOMAIN 154 244 PDZ 2.
FT DOMAIN 486 566 PDZ 3.
FT DOMAIN 600 670 SH3.
FT DOMAIN 768 960 GUANYLATE KINASE.
SQ SEQUENCE 960 AA; 102468 MW; BF87A4262F1B6A5 CRC64;

Query Match 7.5%; Score 379; DB 1; Length 960;
Best Local Similarity 24.8%; Pred. No. 3.2e-13;
Matches 136; Conservative 85; Mismatches 190; Indels 138; Gaps 19;

```

QY 565 IEISKRTGLSLVGSST-LTG---AFIIEVEEGACKDGLMAGDQILEVNGIDL 620
D 41 IOLKNGSLGSLFSTAGCDNDPHIGDTIYITKLISGAAADRLSINDIIVSNVSV 100
QY 621 RKATHEAIVNLROTPOHRLTYRDE-----APKKEEVCDTLTIELQ-KRP 667
D 101 VDPHASAVDALAKAGNVKLVKRRCTATTPAAGSAAAGDASAGPVIETDLVK 160
QY 668 GKGGLSLVGR-----NDTGVFVSDIVKGIADPDGRLLQGOIILV-NG--EDVRN 717
D 161 GKGGLFSTAGIGQNHPIGNDGIVYTKLTDGKQVDRSLIGDLAVRNGSEKLNEN 220
QY 718 ASOBAVALLKCSLGTVLEVGRIKACPFHSERRSPOTSQVSEGLSFTPLSGSTSE 777
D 221 VTHELAVALTKSLTDKVTLLIGK-----TQHLTASGGGGGGLS-----SQOQLSQ 267
QY 778 S---LESSKKNALASEIQLRTVEMKK-GPTDLGISIAGVSPGLDVPFIAMMHT 833
D 268 SOSOLATQSOSQVHQOQHATPYNVSGTGLANSKQTV---VDSF-SIFQAAAANA 322
QY 834 GVAATOTKLRVDRIVTCSTEGMTHTQAVNLKNSGSIEMQVAGGVSV----- 887
D 323 ANASASAVIASNN--TISNTTVTVTATATAS---NDSKLPPLSGANSSISNSNSN 377
QY 888 -----VTGHHOBPASSLSFTGLTSTSI-----FODDLGPP----- 918
D 378 SNSNNINININNNSSSTTATVAATPTAASAAAASPPANFYNNASMPALPV 437
QY 919 -----QCSITLLEKPGDGLG 933
D 438 ESNQTNNSGSPORPQPSRASYNVLAAPGTPRAYSTEDIREPTITIQKPGDLG 497
QY 934 FSIYGVGYPSPHGDPIYKYTFVAKGASEDRGLRKDQIIIVNGSCSLGCVHEEVALIK 993
D 498 FNIYGG---EDGGIYVFLIAGGRADLGSSELKRGDQLSVNNVNLTHAHEEAQALK 553
QY 994 RTKGTVTLM 1002
D 554 TSGGVVTL 562

```

RESULT 11

DLG1_RAT STANDARD; PRT; 911 AA.

AC 062696;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE PRESYNAPTIC PROTEIN SAP97 (SYNAPSE-ASSOCIATED PROTEIN 97) (DISCS, LARGE HOMOLOG 1).

GN DLG1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

OX NCBI_TaxID=10110;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=95198112; PubMed=7891172;

RA Mueller B.M., Kistner U., Voh R.W., Cases-Langhoff C., Becker B., Gundelfinger E.D., Garner C.C.;

RT "Molecular characterization and spatial distribution of SAP97, a novel presynaptic protein homologous to SAP90 and the Drosophila RT discs-large tumor suppressor protein.";

RL J. Neurosci. 15:2354-2366(1995).

CC -1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR SUBUNITS. MAY BE INVOLVED IN SYNAPTOGENESIS. MAY PLAY A ROLE IN CELL ADHESION.

CC -1- SUBCELLULAR LOCATION: IN CULTURED T84 CELLS, IT IS RESTRICTED TO THE CYTOPLASMIC SURFACE OF THE PLASMA MEMBRANES BETWEEN ADJACENT CELLS, BUT NOT AT THE EDGES OF CELLS LACKING CELL-CELL CONTACT.

CC -1- TISSUE SPECIFICITY: PRESYNAPTIC NERVE TERMINI OF EXCITATORY

```

CC SYNAPSES. IN OTHER BRAIN REGIONS, SAP97 IS FOUND IN AND ALONG
CC BUNDLES OF UNMYELINATED AXONS. SAP97 IS NOT RESTRICTED TO THE CNS,
CC BUT IS ALSO PRESENT AT THE BASAL LATERAL MEMBRANE BETWEEN A
CC VARIETY OF EPITHELIAL CELLS.
CC -1- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U14950; AAA79976.1; -
DR HSSP: Q12959; 1PDR.
DR InterPro: IPR000619; -
DR InterPro: IPR001452; -
DR InterPro: IPR001478; -
DR Pfam: PF00625; Guanylate_kin; 1.
DR Pfam: PF00595; PDZ; 3.
DR Pfam: PF00018; SH3; 1.
DR PROSITE: PS00856; GUANYLATE_KINASE_1; 1.
DR PROSITE: PS50052; GUANYLATE_KINASE_2; 1.
DR PROSITE: PS50106; PDZ; 3.
DR PROSITE: PS50002; SH3; 1.
KW SH3 domain; Repeat.
FT DOMAIN 224 310 PDZ 1.
FT DOMAIN 318 404 PDZ 2.
FT DOMAIN 465 545 PDZ 3.
FT DOMAIN 580 650 SH3.
FT DOMAIN 721 911 GUANYLATE_KINASE.
FT DOMAIN 527 530 POLY-A.
SQ SEQUENCE 911 AA; 100570 MW; 18CBBD31DD0CAF8B CRC64;

```

Query Match 7.4%; Score 378.5; DB 1; Length 911;

Best Local Similarity 23.3%; Pred. No. 3.2e-13;

Matches 151; Conservative 91; Mismatches 201; Indels 205; Gaps 21;

```

QY 406 AAYDLSEKKNVQLELPK-----DQGLGIAISEEDTISGVII----- 443
D 48 ALIDIOEYEVTLDDNPKCVDSKQCEPVQGNPMWESSLSAAVTSBLSLPGGLSPVEK 107
QY 444 -----KSLTHGVAARDGRUKVGDQILAUND-----EIVVGPIEKFISLTAKMTV 491
D 108 YRYODEEVLPSERISPOVPNEVLGPELVHSEKSLSEIENYHGFVSHSHSPIKPTF-- 164
QY 492 KLTTHAENPDQSAVPSAAGAASGEKNSOSILAVPOSQPEPESIRMTSRSPAIASD 551
D 165 -----AVPPSPPIYP-----VTPALPPP---AESPVLPSTPQANPPVLYNT 204
QY 552 PA--TCPIIPGCEFTIE-----ISKRTGLSLVGSST-LTG---AFIIEVEEGAA 600
D 205 DSLFETPYVNGTDADEYEIEETLTERGNSGLGFSIAGGTDPNDPHIDDSIFTKIITGAA 264
QY 601 CKDQRLMAGDQILEVNGIDLRKATHDEAIVNLROTPOHRLTYRDSAPYKKEEVCOTLT 660
D 265 AQDQRLVNDGILRVNEADVADVTHSKAVEALKAGSIVRLVYRRKAFKRNHE----- 318
QY 661 IELKKPKGKGLSLVGR-----NDTGVFVSDIVKGIADPDGRLLQGOIILVNGED 714
D 319 IKLKGPF-KGLGFSIAGVGQNHPIGNSIYVTKIIEGAAHKRGKQIADKLLAVNSVC 377
QY 715 VRNASQBAVALLKCSLGTVLEVGRIKACPFHSERRSPOTSQVSEGLSFTPLSGSS 774
D 378 LEEVTHAEAVATLAKTSDFVYLKAAK-----PTSMYINDG-----YAPDPITNS 421
QY 775 TSELESSKKNALASEIQLRTVEMKKGPTDLGISIAGVSPGLDVPFIAMMHTG 834

```

```

DB 422 SSGVDNHWSPS-----SYLGQFPASPARSPIS 450
QY 835 VAAQGTOKLRVDRIVTICSTEGMTHTQAVNLKNASSIEMOVYAGGVSVTGHQOE 894
DB 451 KAV-----LGDEBIT-----RE 462
QY 895 PASSSLFTGLNSTSIFQDDLPPOCKSITLERGPDGLGFSYGVGSPHGLPIYKVTY 954
DB 463 P-----RKVYLHSGSTGLGPNIVG---EDGE-GIFISFT 493
QY 955 FAKGAASEDGRKRDQDIIAVNGOSLEGVTHEEVAAILKRTKGTITLM 1002
DB 494 LAGGPADLSELGRKDRITISVNSVDLRAASHEDQAAALKNAGQAVTIV 541

RESULT 12
DGL1_HUMAN
ID DGL1_HUMAN STANDARD; PRT; 904 AA.
AC Q12959; Q12958;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE PRESYNAPTIC PROTEIN SAP97 (SYNAPSE-ASSOCIATED PROTEIN 97) (DISCS,
DE LARGE HOMOLOG 1).
GN DGL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=95024052; Pubmed=7937897;
RA Lue R.A., Marfatia S.M., Branton D., Chishti A.H.;
RT "Cloning and characterization of hdl9: the human homologue of the
RT Drosophila discs large tumor suppressor binds to protein 4.1.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:9818-9822(1994).
RN (2)
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 460-555.
RX MEDLINE=96338231; Pubmed=8757139;
RA Cabral J.H., Petosa C., Sutcliffe M.J., Raza S., Byron O., Poy F.,
RA Marfatia S.M., Chishti A.H., Liddington R.C.;
RT "Crystal structure of a PDZ domain.";
RL Nature 382:649-652(1996).
CC -1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR
CC SUBUNITS. ASSOCIATES WITH PROTEIN 4.1.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: CONTAINS 3 PDZ/DRH DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U13897; AAA50599.1; -
CC EMBL: U13896; AAA50598.1; -
CC PDB: 1PDR; 23-JUL-97.
CC MIM: 601014; -
CC InterPro: IPR000619; -
CC InterPro: IPR001452; -
CC InterPro: IPR001478; -
CC Pfam: PF00625; Guanylate_kin; 1.
CC Pfam: PF00595; PDZ; 3.
CC Pfam: PF00018; SH3; 1.
CC PROSITE: PS00856; GUANYLATE_KINASE_1; 1.
CC PROSITE: PS0052; GUANYLATE_KINASE_2; 1.
CC PROSITE: PS0106; PDZ; 3.

```

```

DR PROSITE; PS00002; SH3; 1.
KW SH3 domain; Repeat; Alternative splicing; 3D-structure.
FT DOMAIN 224 310 PDZ 1.
FT DOMAIN 319 405 PDZ 2.
FT DOMAIN 466 546 PDZ 3.
FT DOMAIN 581 651 SH3.
FT DOMAIN 714 904 GUANYLATE KINASE.
FT VARSPIC 669 680 EIPDDMGSKGLK -> OSFNDKRNKRLFRKPPFYKNDQS
FT SEQUENCE 904 AA; 100354 MM; B78798D6B0920D4 CRC64;

Query Match 7.4%; Score 377; DB 1; Length 904;
Best local similarity 29.9%; Pred. No. 3 Be-13;
Matches 119; Conservatve 58; Mismatches 129; Indels 92; Gaps 14;

QY 634 QTPQRYRLTLRYDEAPYKEEVCDTLTTELQKPRGGLISYGR-----NDTGVFVS 687
DB 208 ETPTYVNGT---DAVEYEI---TLE---RNGSLGFSINGTDNPHIGDSSIFIT 286
QY 688 DIYKGGIADPDGRLIOGDQILVNGEDVRYNASQEAVALKCSLGYTTEVGRKAGPRH 747
DB 257 KIITGGAAADGRLRYNDCLIOVNEVDYRDVTSKAVEALKKSGSYRLVYKR----- 309
QY 748 SEHRPSQTSQVSEGLSSEFFPLSGSTSESSSKNALASDIQGLRTVEKKKPTDS 807
DB 310 --RKP-----VSE-----KIMEIKLIKGP-KG 328
QY 808 LGISIAAGVSP--LGDVPIFIAMHPTGYVAQOTOKLRVDRIVTICSTEGMTHTQAV 865
DB 329 LGFSIAGVGNGHPIPGNSIYVTKIIEGGAHDKLQIGDKLLAVNNVCLLEVTHEAV 388
QY 866 NLKNASGSIEMOVYAGGVSV-----VTGHQEPAS---SSLSFTLT----- 906
DB 389 TALKNTSDFYLLKVAAPTSYKMDGYAPDITNSSQPDVNHVSPSFLQTPASPARYS 448
QY 907 --STSIQDDLPPOCKSITLERGPDGLGFSYGVGSPHGLPIYKVTYFARGAASEDG 964
DB 449 PVKAVAYGDEETREPRKVVYLHSGSTGLGPNIVG---EDGE-GIFISFLAGPADLSG 504
QY 965 RLKRGDQIIAVNGOSLEGVTHEEVAAILKRTKGTITLM 1002
DB 505 ELKRGDRITISVNSVDLRAASHEDQAAALKNAGQAVTIV 542

RESULT 13
IL16_HUMAN
ID IL16_HUMAN STANDARD; PRT; 631 AA.
AC Q14005; Q16435; Q9UP18;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE INTERLEUKIN-16 PRECURSOR (IL-16) (LYMPHOCYTE CHEMOKINE/TRACTANT FACTOR)
DE (LCEP).
GN IL16.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=peripheral blood;
RA Kornfeld H.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=99145586; Pubmed=9990060;
RA Bannert N., Avots A., Baler M., Serfling E., Kurth R.;
RT "GA-binding protein factors, in concert with the coactivator CREB
RT binding protein/p300, control the induction of the Interleukin 16
RT promoter in T lymphocytes.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:1541-1546(1999).
RN (3)

```

RP SEQUENCE OF 241-631 FROM N.A.
 RA MEDLINE-96211141; PubMed-8609984;
 RA Bannert N., Baier M., Werner A., Kurth R.;
 RT "Interleukin-16 or not?";
 RL Nature 381:30-30(1996).
 RN [4]
 RP SEQUENCE OF 502-631 FROM N.A.
 RC Tissue-Peripheral blood;
 RX MEDLINE-94255480; PubMed-7910967;
 RA Crikshank W.W., Center D.M., Nisar N., Wu M., Natke B.C.,
 RA Theodore A.C., Kornfeld H.;
 RT "Molecular and functional analysis of a lymphocyte chemoattractant
 RT factor: association of biologic function with CD4 expression.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:5109-5113(1994).
 RN [5]
 RP SEQUENCE OF 502-631 FROM N.A.
 RA Du Y., Du G.X., Hou L.H., Wang H.T.;
 RT "CDNA sequence of Interleukin-16 cloned from human PBLS.";
 RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: STIMULATES A MIGRATORY RESPONSE IN CD4+ LYMPHOCYTES,
 CC MONOCYTES, AND EOSINOPHILS. ALSO INDUCES T-LYMPHOCYTE EXPRESSION
 CC OF INTERLEUKIN 2 RECEPTOR. LIGAND FOR CD4.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- SIMILARITY: CONTAINS 2 PDZ/DHR DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; M90391; AAD04636.1; -
 DR EMBL; AF077011; AAD15990.1; -
 DR EMBL; S81601; AAB36371.1; -
 DR EMBL; AF053412; AAC12732.1; -
 DR MIM; 603035; -
 DR HSSP; Q12959; 1PDR.
 DR InterPro: IPR001478; -
 DR Pfam: PF00595; PDZ 1.
 DR PROSITE: PS50106; PDZ 2.
 KW Cytokine; Chemotaxis.
 FT PROPEP 1 ?
 FT CHAIN 1 ?
 FT DOMAIN 411 631 INTERLEUKIN-16.
 FT DOMAIN 533 618 PDZ 1.
 FT DOMAIN 104 104 PDZ 2.
 FT CONFLICT 233 233 E -> D (IN REF. 2).
 FT CONFLICT 241 243 F -> L (IN REF. 2).
 FT CONFLICT 319 319 LRL -> PRE (IN REF. 3).
 FT CONFLICT 319 319 E -> A (IN REF. 1).
 SO SEQUENCE 631 AA; 66694 MW; 2A9CED70CC8D4AC CRC64;

Query Match 5.68; Score 283; DB 1; Length 631;
 Best Local Similarity 22.0%; Pred. No. 2.8e-08;
 Matches 171; Conservative 106; Mismatches 306; Indels 194; Gaps 31;

QY 256 SODVKEDEFGSMKNIERYGTGLGELHMELEKSGGLGSLAGNDRSMSTFIYGI 315
 DB 4 SFDITAEPPWARISSCINLNPSINSEH-----GHPLOPNASLNEEGTQG----- 51
 QY 316 DPNGAAGRGRLQIADLELLEITNGQLYGRSHQNA-SSTIKCAPS-----KVIIFIR 366
 DB 52 HPDGRPR-----LDITANGT---PKYKSAOSSIVYKGPYPAPKPRFROSLKGLR 99
 QY 367 NKDAVNAQVCPGNAVEPLPSNSENLQKREPTVTSDAVAIVSSFKNVQHLLEPKDQG 426
 DB 100 NRASPRGLPPLALSTGAPAPASREHL-GSHIRASSSSSIRORISSFETGSSQLP-DKG 157
 QY 427 GLGIAISEDLISGVIITSLTEHGVAATDGRKAGDQILAVDELIVGPIEKFISLTKT 486
 DB 158 AQRSLQPS---SGEAPKPLGNH-----EGR-----FSGLT-- 186
 QY 487 AKMTKLTIHENPDSQAVPASAASEKKNSSQSLWPOSGSPESINITSSTPA 546
 DB 187 -----GRGA-----PTLVPO---QPRV--LSSGSPA 210
 QY 547 IFASDPATPIPGCETTYEISKRTGLGLSIVGSDPLTAFIIEYERGAACKDGL 606
 DB 211 SEARPPGVSESP-----PGRQPNQKTFPPPPDPLRLSLQARESGGP----- 254
 QY 607 WAGDQILEVNGIDLRKATHEAIVNLRQTPQVRVLTLYRDEAPYKKEEYCDTLTELQK 666
 DB 255 -----VLKMPQORAR-----SPPLRSQSCETKL--IDEX 282
 QY 667 PGKGLGIVGKRNQDGVFVSDIVVKGADDPGRRLQGDQILVLNGEDVVRNAPAVAL 726
 DB 283 TSKLXIS-----SOVSSAVMKSLLCLPSSISCAQTPCIPKGAAPTSSNEDSA- 333
 QY 727 LKCSIGTVTLEVGRKAGPFH---SERP-----SQTQVSEGLSISPTFPLSGSTSES 778
 DB 334 -NGSAETSAIDTG-----FSLNISELREYEGLEAKNEDDGDHSLQ---SGGSVSL 383
 QY 779 LESSSK-----NALSEIOGLTVEKKKPTSLGISTINGVSGPLGDVPPIA 828
 DB 384 LSSEELKRLIEVKYLDATLQDLQIHVTLLHKEGAGLESLAG--GAULENKVTVH 441
 QY 829 MMHPTGVAQOTOKLRVGRIVTICSTEGMTHTQAVMLKNASGSIEMQVAVAGDVSV 888
 DB 442 RVFPRGLASQESTIOKGNHEVLINSLKLGTHHALILNQARPRQAVYTRKLTPEA 501
 QY 889 TGHQEPASSLSIFGLTSTSFDDDLPPPOCKSTTLERGDGLGFSIVGYSHGCLP 948
 DB 502 MPDLNSTSDSAASASMAADSV--ESTAEAVYCVTTLKMSAGLGFSLGKGSJHGDKP 559

QY 949 IYKTVFAKGAASEDGR-IAKRODQIIANQOSLEGVTEENVAIKR-TKGVTLMV 1003
 DB 560 LTINRIF-KGAASEQSEVYQFODELQGLGTRMOSLTFEAMNITKALPDGPVITI 615

RESULT 14
 ZOI_HUMAN
 ID ZOI_HUMAN STANDARD; PRT: 1736 AA.
 AC 007157;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE TIGHT JUNCTION PROTEIN ZO-1 (ZONA OCCUDENS 1 PROTEIN) (TIGHT JUNCTION
 DE PROTEIN 1).
 GN TUP1 OR ZO1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Liver;
 RX MEDLINE-93361541; PubMed-8395056;
 RA Willott E., Balda M.S., Fanning A.S., Jameson B., van Itallie C.,
 RA Anderson J.M.;
 RT "The tight junction protein ZO-1 is homologous to the Drosophila
 RT discs-large tumor suppressor protein of septate junctions.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:7834-7838(1993).
 CC -1- FUNCTION: THE N-TERMINAL MAY BE INVOLVED IN TRANSDUCING A SIGNAL
 CC REQUIRED FOR TIGHT JUNCTION ASSEMBLY, WHILE THE C-TERMINAL MAY
 CC HAVE SPECIFIC PROPERTIES OF TIGHT JUNCTIONS. THE ALPHA DOMAIN
 CC MIGHT BE INVOLVED IN STABILIZING JUNCTIONS.
 CC -1- SUBUNIT: INTERACTS WITH OCCUDIN, CLAUDINS AND ZO-3.
 CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE, CYTOPLASMIC SIDE.
 CC MOVEMENT OF ZO-1 FROM THE CYTOPLASM TO MEMBRANE IS AN EARLY EVENT
 CC OCCURRING CONCURRENTLY WITH CELL-CELL CONTACT.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: THE ALPHA-CONTAINING ISOFORM IS FOUND IN MOST

Query	Best Local Similarity	4.8% Score 243; DB 1; Length 1736;
Matches 100; Conservative 68; Mismatches 149; Indels 182; Gaps 15;		
QY 651 KEEVYCDTLTIELQPKPKGLSLIVGRN-----DTGVVSDIVKGLADPDGRLIQ 703	100	100
DB 2 EETATWECHTYTLHNAPEFGFGLIALSGGDNDNHPGSGETSIYIDVLKGGPA--EGQLDE 59.	100	100
QY 704 GDQIILVNGEDVRNAAQ--AAVALIKSLGT-----VLEVGRIKAGFFHSERRPS 753	100	100
DB 60 NDRVAMVNGVSDMNEHAFAVQOLKRKSGKNAKITIRRKKKVQIPVSRDPEVSDNEEDS 119	100	100
QY 754 QTSQVSEBSLSFTPLPLG-----STSSLESSSKKNKLASE--IQ 793	100	100
DB 120 YDEETHD-----PRGGRSGVNVNRSEKTIWPRDRSASRERSLSPRSDDRVASVSSQPAK 171	100	100
QY 794 GLRTYEMKKGPTDLSIGISIVAGVSPGLDVPFIAMHPTGVAAGTOKLRVGDIVTICG 853	100	100
DB 172 PTKVTLVRSKRNEEGLALAS-----HIFVKEIQSDSLAANDGNIGBEDVYLKING 222	100	100
QY 854 TSTEGMTHTQAVNLKNAAGSIEMQVAG-----GDVSV 887	100	100
DB 223 TVTEWMSLTDAKTLIERKSKLKMVYQDERATLLNVPLSDSISANASERDDISEIQS 282	100	100
QY 888 VTGHH-----QEPASSL-----SFGGLNST 908	100	100
DB 283 LASDHSGRSHDRPPRRSRSPDRQSEPSDHSRHSPOQPSNGSLRSRDEERISKGAVST 342	100	100

```

OY 909 SI-FODDGLPGCKSITLERGP-----|||-----929
Db 343 PVKHAADHTPTVEVETVYERNEKOTPSLPEPKRPVYAQVGNOMMITYLSVHNAVSTIQLMK 402
OY 930 -----DELGFSIVGGYGSPPHGLDPIYKVTYFAKGAASEDRGLKRGDQII 973
Db 403 MGLRPSMKLVKFRKRGDSVGLRLAGC-----NDVGIFVAGVLEDSPAKNEG-LESGDQIL 456
OY 974 AVNGOSLEGVTHEEAVAIL 992
Db 457 RVNNVDEFTNIREEAVLEFL 475

RESULT 15
Z03_HUMAN STANDARD; PRT; 933 AA.
AC 095049;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TIGHT JUNCTION PROTEIN ZO-3 (ZONA OCCUDENS 3 PROTEIN) (TIGHT JUNCTION
DE PROTEIN 3).
DE TP3 OR Z03.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamerich J.E., McGready P.M., Skowronski E., Vysvanathan V.,
RA Burkhardt-Schultz K., Brower A., Gordon L., Dias J., Ramirez M.,
RA Stillegen S., Phan H., Velasco N., Do L., Regala W., Terry A.,
RA Ganes J., Danganan L., Ertler A., Christensen M., Georgescu A.,
RA Avila J., Liu S., Attk C., Andreise T., Trankheilm M.,
RA Amico-Keller G., Coefield J., Duarte S., Lucas S., Bruce R.,
RA Thomas P., Quan G., Kronmiller B., Arellano A., Sanders C., Ow D.,
RA Nolan M., Trong S., Kobayashi A., Olsen A.S., Carrano A.V.;
RA "Sequence analysis of a 3.5 kb contig in human 19p13.3 containing a
RT serine protease gene cluster."
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1 SUBUNIT: INTERACTS WITH OCCUDLIN, CLAUDINS AND ZO-1 (BY
CC SIMILARITY).
CC -1 SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
CC -1 SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1 SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
CC -1 SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AC005954; AAC7274.1; ALT_INIT.
DR HSSP: P31016; 18FE.
DR InterPro: IPR000619; -
DR InterPro: IPR001452; -
DR InterPro: IPR001478; -
DR Pfam: PF00595; PDZ 3.
DR Pfam: PF00625; Guanylate_kin; 1.
DR PROSITE: PS00856; GUANYLATE_KINASE_1; FALSE_NEG.
DR PROSITE: PS50052; GUANYLATE_KINASE_2; 1.
DR PROSITE: PS50002; SH3; 1.
DR PROSITE: PS50106; PDZ; 3.
DR Tight junction; SH3 domain; Repeat; Membrane.
FT DOMAIN 11 93 PDZ 1.
FT DOMAIN 195 272 PDZ 2.
FT DOMAIN 394 460 PDZ 3.
FT DOMAIN 489 563 SH3.
FT DOMAIN 675 775 GUANYLATE KINASE.

```

SEQUENCE 933 AA; 102800 MM; 680298CFD0615B47 CRC64;

Query Match 4.88; Score 242; DB 1; Length 933;

Best Local Similarity 22.88; Pred. No. 8.2e-06; Matches 121; Conservative 62; Mismatches 155; Indels 192; Gaps 20;

QY 564 TIEISKG-RTGLGSIYVGSPTLGAFLIHEVEGAACKDGLMAGDQILEVNGIDLK 622
 Db 11 TATLSKDPRRGFGTALISGGRDRPGSMVSDVYVPGPA-EGRLQTGDHIYVNGVSMEN 68
 QY 623 ATHDEAINVLR-----QTPOVRLTLXR-----DEAPYKEEYV----- 655
 Db 69 ATSAFAIQILKTCRKMANITVAPRRIRHLPATKASPPSSPGRODSDDEDDGPQREVEVDGR 128
 QY 656 -----CDTLTI-----EL 663
 Db 129 GYDDSSSGSGSRWDERRRRPRGRAGSHGRSPGSGSEANGALALVSGFKRLPRQDV 188
 QY 664 QKKPGKGLGSIYVGRND-----TGVFVSDIVKGIADPDGRLIQDQILLVNGED 714
 Db 189 QMKPVK---SVLVKRRDSEFEGVKGQIFIKHITDGLAARHRLQEGDLILQINGVS 244
 QY 715 VRNASQEAVALLKCSICTVTLVGR----- 740
 Db 245 SONSLNDTRRLIEKSEGLSLVLRDGRQFLVNIIPRAVSDSDSSPLEEGVTMADEMSSP 304
 QY 741 -----IKAGPFHSERRP--SOTSQV-----SEGLSSTFEPPLSGS 773
 Db 305 PADISDLASELSQAPPSHIIPPRHAKRSPASQTDSPVESPRLRRESSVDSRTISEPDE 364
 QY 774 STSESLESSS-----KKNALASEIQG---LRTYEMKKGPTDSLGISTAGVGSPLGDVP 824
 Db 365 QRSLEPRESSYDIYRVPSQSMEDRGYSPDTRVYVFLKG--KSIGLRLAGC-----NDVG 417
 QY 825 IFIMMHPTGYAAQTKRVGDRTVTCPTSTEGHTQAVNLKNAGSTIEMOVVAG-- 882
 Db 418 IFVSGVQ-AGSPADGGQIQEGDQILQVNDVPFQNLTRREAVQFLGLPPGGEEMELVYORK 476
 QY 883 -----GDVSVYTGHHO--EPASSLSFTGLTSTSIQO--DOLGP 917
 Db 477 QDIFKMKVQSRVGDSTFIRTHFELEP--SPPSGLGFTRGDYFHVLDITLHP 524

Search completed: July 12, 2001, 14:43:25
 Job time: 218 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2001, 14:43:01 ; Search time 64.35 Seconds
(Without alignments)
2066.302 Million cell updates/sec

Title: US-09-502-698-2

Perfect score: 5085
Sequence: 1 MGNVSKSEFERTINAKGN.....EEAVALLKRTKGVTLWLIS 1005

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMBL_16:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp_invertebrate:*
7: sp.mammal:*
8: sp.mhc:*
9: sp.phage:*
10: sp.plant:*
11: sp.prodent:*
12: sp.unclassified:*
13: sp.vertebrate:*
14: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4839	95.2	2042	4	075970
2	4413.5	86.8	2054	11	055164
3	4314.5	84.8	2055	11	0921K3
4	2297.5	45.2	526	11	008783
5	2252	44.3	453	4	043798
6	1496.5	29.4	1134	4	09H3N9
7	1469.5	28.9	612	11	070471
8	1223.5	24.1	1552	4	060833
9	1198.5	23.6	1582	4	043742
10	1063.5	20.9	1524	4	015249
11	814	16.0	2208	5	009515
12	589	11.6	728	11	070263
13	564.5	11.1	665	5	023823
14	553.5	10.9	674	5	09NBV3
15	551.5	10.8	674	5	024008
16	534	10.5	1012	5	021075
17	507.5	10.0	2460	11	064512
18	507	10.0	2484	6	028006
19	506	10.0	1551	4	014160

20	453.5	8.9	1256	4	075085	075085 homo sapien
21	443	8.7	2766	11	09QZR8	09QZR8 rattus norv
22	441	8.7	1150	4	09HBC4	09HBC4 homo sapien
23	432	8.5	1125	4	09H2V6	09H2V6 homo sapien
24	431.5	8.5	1179	11	09JRW1	09JRW1 rattus norv
25	429.5	8.4	1126	11	09ECQ9	09ECQ9 mus musculu
26	425	8.4	874	4	09HCD8	09HCD8 homo sapien
27	424	8.3	871	5	09X235	09X235 drosophila
28	422	8.3	871	5	09NB04	09NB04 drosophila
29	419.5	8.2	1756	5	09VBE4	09VBE4 drosophila
30	403	7.9	1278	4	060434	060434 homo sapien
31	399.5	7.9	1131	5	08165	08165 caenorhabdi
32	396.5	7.8	1112	11	09WQ01	09WQ01 mus musculu
33	395	7.8	1277	11	088382	088382 rattus norv
34	394	7.7	1114	11	09R271	09R271 rattus norv
35	392	7.7	1455	4	060510	060510 homo sapien
36	385	7.6	960	5	09VY25	09VY25 drosophila
37	385	7.6	1171	11	054893	054893 mus musculu
38	381	7.5	950	11	09WU36	09WU36 rattus norv
39	377.5	7.4	822	11	088961	088961 rattus norv
40	377.5	7.4	1043	11	09WTW1	09WTW1 rattus norv
41	376	7.4	1322	11	09Q2P6	09Q2P6 mus musculu
42	370	7.3	1027	4	09H1X7	09H1X7 homo sapien
43	368	7.2	927	11	062402	062402 mus musculu
44	357	7.0	1112	11	097879	097879 rattus norv
45	352	6.9	1266	4	09NYE6	09NYE6 homo sapien

ALIGNMENTS

RESULT 1

ID	075970	PRELIMINARY:	PRT:	2042 AA.
AC	075970;			
DT	01-NOV-1998 (TREMBLrel. 08, Created)			
DT	01-NOV-1998 (TREMBLrel. 08, Last sequence update)			
DT	01-MAR-2001 (TREMBLrel. 16, Last annotation update)			
DE	MULTI PDZ DOMAIN PROTEIN MUPPL.			
GN	MUPPL.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=BRIN;			
RA	Eng L., Kravitsky G., Clapham D.E.;			
RT	"Human homolog of Muppl protein."			
RL	Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AF093419; AAC61870.1; -.			
DR	HSSP; Q12959; 1PDR.			
DR	InterPro; IPR001478; -.			
DR	Pfam; PF00595; PDZ; 13.			
DR	SMART; SM00228; PDZ; 1.			
DR	SEQUENCE 2042 AA; 218595 MW; FOEIDDD29749A762 CRC64;			

Query Match 95.2% Score 4839; DB 4; Length 2042;
Best Local Similarity 90.2% Pred. No. 9.3e-284;
Matches 970; Conservative 4; Mismatches 3; Indels 98; Gaps 2;

QY	1	MGNVSKSEFERTINAKGNSSLGKMTVSNKKGGLGMYRSIIHGAISMDGIAIGDCTL 60
DB	996	MGNVSKSEFERTINAKGNSSLGKMTVSNKKGGLGMYRSIIHGAISMDGIAIGDCTL 1055
QY	61	STNESTISVTNAQARAMLRHSISGPDIKITYVAEHEEERKISLGGQSGRVMLDIFS 120
DB	1056	STNESTISVTNAQARAMLRHSISGPDIKITYVAEHEEERKISLGGQSGRVMLDIFS 1115
QY	121	SYTGRIPELPREGESESELQNTAVSNMQPRVELMRPSSKLSIYVGRGMSR 180
DB	1116	SYTGRIPELPREGESESELQNTAVSNMQPRVELMRPSSKLSIYVGRGMSR 1175

```

OY 181 LSNQVNRGIFIKHVLKEDSPAGKNGTLKPGDRIVE-----215
DB 1176 LSNQVNRGIFIKHVLKEDSPAGKNGTLKPGDRIVEVDGMOLRQASHQAVAIRKAGNPV 1235
OY 216 -----APQSESEPEKAPLC 230
DB 1236 VFMVQSIINRPRKSPPLSLHLNLYPKYNSFSTNPFADSLQIMDKAPQSQSEPEKAPLC 1295
OY 231 SVPPPSAFAEMGSDHTQSSASKISODLQKDEDFGYSWKNIIRRYCTLTGELHMIIELEK 290
DB 1296 SVPPPSAFAEMGSDHTQSSASKISODVDKDEDFGYSWKNIIRRYCTLTGELHMIIELEK 1335
OY 291 GHSGSLGSLAGNDRSMYSFIVGIDPNGAAGKGRLOIADDELLEINGQILLYGRSHONAS 350
DB 1356 GHSGSLGSLAGNDRSMYSFIVGIDPNGAAGKGRLOIADDELLEINGQILLYGRSHONAS 1415
OY 351 SIICAPSKVKIIFIRNKDAVNQMAVCPGNAVEPLPSNSENLOKKEPEPTVTSDAAYDL 410
DB 1416 SIICAPSKVKIIFIRNKDAVNQMAVCPGNAVEPLPSNSENLOKKEPEPTVTSDAAYDL 1475
OY 411 SSFKNVHLELPKQGGSLGIAISEEDTLGVIITKSLTEHGVAATDGRKXGDOILLAVDDE 470
DB 1476 SSFKNVHLELPKQGGSLGIAISEEDTLGVIITKSLTEHGVAATDGRKXGDOILLAVDDE 1535
OY 471 IVVGYPIEFISLTKAKMTYVKTILHAENPDQAVPSAGAAGEKKNSOSLWVPOSGS 530
DB 1536 IVVGYPIEFISLTKAKMTYVKTILHAENPDQAVPSAGAAGEKKNSOSLWVPOSGS 1595
OY 531 PEPSINTSRSSSTPAIFASDPATCPIIPGCEETIEISKRTGLGSLIVGSDTLGAFI 590
DB 1596 PEPSINTSRSSSTPAIFASDPATCPIIPGCEETIEISKRTGLGSLIVGSDTLGAFI 1655
OY 591 IHEVYEEGAAKCDKRLMAGQOILEVNGIDLRKATHDEAINVLRQTPQVRLLTLRDEAPY 650
DB 1656 IHEVYEEGAAKCDKRLMAGQOILEVNGIDLRKATHDEAINVLRQTPQVRLLTLRDEAPY 1715
OY 651 KEEBVCDTLTLELQKPKGKGLGSLIVGKRNDYGVFVSDIYKGIADPDGRLIOGDOILLV 710
DB 1716 KEEBVCDTLTLELQKPKGKGLGSLIVGKRNDYGVFVSDIYKGIADPDGRLIOGDOILLV 1775
OY 711 NGEVNRASQDAVAALLKCSLGYTLELVGRYKACPFHSERRPSOTSOVSEGLSSFFPPL 770
DB 1776 NGEVNRASQDAVAALLKCSLGYTLELVGRYKACPFHSERRPSOTSOVSEGLSSFFPPL 1808
OY 771 SGSTSESLSSSKKNALASEIOGLRTVEKKKCPDLSIGISIAGVGSPLGDVPIFIAMM 830
DB 1809 -GSSTSESLSSSKKNALASEIOGLRTVEKKKCPDLSIGISIAGVGSPLGDVPIFIAMM 1867
OY 831 HPFGVAAGTOKLRVGRIVITICGTSTEGMTHTQAVNLLKNAAGSIEQVYVAGDVSVAVTG 890
DB 1868 HPFGVAAGTOKLRVGRIVITICGTSTEGMTHTQAVNLLKNAAGSIEQVYVAGDVSVAVTG 1927
OY 891 HHQEPASSISLFTGLTSTSTFODDLGPPROCKSTLLEKPGGLGSLIVGGSPPGDLPIY 950
DB 1928 HHQEPASSISLFTGLTSTSTFODDLGPPROCKSTLLEKPGGLGSLIVGGSPPGDLPIY 1987
OY 951 VKTVEFAGASSEDGRILKRGDOIIIVANGQSLEGYTHEEVAVALIKRTKGTIVLWVLS 1005
DB 1988 VKTVEFAGASSEDGRILKRGDOIIIVANGQSLEGYTHEEVAVALIKRTKGTIVLWVLS 2042

```

```

RESULT 2
O55164 ID 055164 PRELIMINARY; PRT: 2054 AA.
AC 055164:
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DE 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE MULTI PDZ DOMAIN PROTEIN 1.
GN MUPPL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98196865; PubMed=9537516;
RA Ulmer C., Schuck K., Flage A., Lubbert H.,
RT "Cloning and characterization of MUPPL, a novel PDZ domain protein."
RL FEBS Lett. 424:63-68(1998).
DR EMBL: AJ001320; CA004681.1;
DR HSSP: Q12959; 1PDR
DR InterPro: IPR001478;
DR Pfam: PF00595; PDZ; 13.
DR SMART: SM00228; PDZ; 1.
SO SEQUENCE 2054 AA; 218590 MW; 44BD3F42B801F78F CRC64;

```

```

Query Match 86.8%; Score 4413.5; DB 11; Length 2054;
Best Local Similarity 81.5%; Pred. No. 5.1e-258;
Matches 875; Conservative 64; Mismatches 62; Indels 73; Gaps 2;

```

```

OY 2 LQVSKSEPFRTINAKNSLIGMTVSANKDGLMIVRSIIHGALSRDRIAGDCLIS 61
DB 984 LQVSKSEPFRTINAKNSLIGMTVSANKDGLMIVRSIIHGALSRDRIAGDCLIS 1043
OY 62 INESTISVTNAQARALRRHSLIGPDIKITYVPAEHLKEFKISLGOOSGRVALDIFSS 121
DB 1044 INESTISVTNAQARALRRHSLIGPDIKITYVPAEHLKEFRVSVFGQAGIAMLDFSS 1103
OY 122 YTGADIDELPEBEEGESEELONTAVSNNOPRVELMRPESKLSGISIVGKMGSR 181
DB 1104 YTGADIDELPEBEEGESEELONTAVSNNOPRVELMRPESKLSGISIVGKMGSR 1163
OY 182 SNGEVMGIFIKHVLKEDSPAGKNGTLKPGDRIVE-----215
DB 1164 SNGEVMGIFIKHVLKEDSPAGKNGTLKPGDRIVEVDGMOLRQASHQAVAIRKAGNPV 1223
OY 216 -----APQSESEPEKAPLC 231
DB 1224 FMVQSIINRPRKSPPLSLHLNLYPKYNSFSTNPFADSLQIMDKAPQSQSEPEKAPLC 1283
OY 232 VPPPPSAFAEMGSDHTQSSASKISODVDKDEDFGYSWKNIIRRYCTLTGELHMIIELEK 291
DB 1284 VPPPPSAFAEMGSDHTQSSASKISODVDKDEDFGYSWKNIIRRYCTLTGELHMIIELEK 1343
OY 292 HSGSLGSLAGNDRSMYSFIVGIDPNGAAGKGRLOIADDELLEINGQILLYGRSHONAS 351
DB 1344 HSGSLGSLAGNDRSMYSFIVGIDPNGAAGKGRLOIADDELLEINGQILLYGRSHONAS 1403
OY 352 SIICAPSKVKIIFIRNKDAVNQMAVCPGNAVEPLPSNSENLOKKEPEPTVTSDAAYDL 411
DB 1404 SIICAPSKVKIIFIRNKDAVNQMAVCPGNAVEPLPSNSENLOKKEPEPTVTSDAAYDL 1463
OY 412 SSFKNVHLELPKQGGSLGIAISEEDTLGVIITKSLTEHGVAATDGRKXGDOILLAVDDE 471
DB 1464 SSFKNVHLELPKQGGSLGIAISEEDTLGVIITKSLTEHGVAATDGRKXGDOILLAVDDE 1523
OY 472 IVVGYPIEFISLTKAKMTYVKTILHAENPDQAVPSAGAAGEKKNSOSLWVPOSGS 531
DB 1524 IVVGYPIEFISLTKAKMTYVKTILHAENPDQAVPSAGAAGEKKNSOSLWVPOSGS 1580
OY 531 EPESINTSRSSSTPAIFASDPATCPIIPGCEETIEISKRTGLGSLIVGSDTLGAFI 591
DB 1581 EPESINTSRSSSTPAIFASDPATCPIIPGCEETIEISKRTGLGSLIVGSDTLGAFI 1640
OY 592 HEVYEEGAAKCDKRLMAGQOILEVNGIDLRKATHDEAINVLRQTPQVRLLTLRDEAPY 651
DB 1641 HEVYEEGAAKCDKRLMAGQOILEVNGIDLRKATHDEAINVLRQTPQVRLLTLRDEAPY 1700
OY 652 EEBVCDTLTLELQKPKGKGLGSLIVGKRNDYGVFVSDIYKGIADPDGRLIOGDOILLV 711
DB 1701 EEBVCDTLTLELQKPKGKGLGSLIVGKRNDYGVFVSDIYKGIADPDGRLIOGDOILLV 1760

```

Qy	712	GEVRNNAEOEVNALLKCSGTYVTEVGRKAKPFHSERBPOTSOVSQSGISLSTFPFLS	771
Db	1761	GEDVRNNAEOEVNALLKCSGTYVTEVGRKAKPFHSERBPOTSOVSQSGISLSTFPFLS	1820
Qy	772	GSSTSESLSSSKKNALASEIOGLRTVEYMKKGPDSIGISAGVSPGLDVPITIAMH	831
Db	1821	GIHSTSESSSSAKKNALASEIOGLRTVEIKKGPALDLSIAGVSPGLDVPITIAMH	1880
Qy	832	PTGVAATOKLRNDRLVYTCGTSTEGMHTQANVLKNASSIEHQVYVAGSDVSVMGH	891
Db	1881	PNGVAATOKLRNGDRILVYTCGTSTEGMHTQANVLKNASSIEHQVYVAGSDVSVMGH	1940
Qy	952	HOEPAASSLSTGTLTSTSIIFODDLGPPOCKSITLERGPGDLFSIYGVGSPHGDLPITY	951
Db	1941	QOELANPCLAFGTGLTSTIIPDDLGPPQSKTILDRGPDGLFSIYGVGSPHGDLPITY	2000
Qy	952	KTVPAKKAASDGLKLRGGDIIIVANGQSLSEGYVHEBAVALIKRTKGTVYIAML5	1005
Db	2001	KTVPAKKAALADGKLKRGDIIIVANGQSLSEGYVHEBAVALIKRTKGTVYIAML5	2054

Q9Z1K3	3		
ID	Q9Z1K3	PRELIMINARY;	PRT; 2055 AA.
AC	Q9Z1K3;		
DT	01-MAY-1999	(TREMBLrel. 10, Created)	
DT	01-MAY-1999	(TREMBLrel. 10, Last sequence update)	
DT	01-MAR-2001	(TREMBLrel. 16, Last annotation update)	
DE	MULTIPLE PDZ DOMAIN PROTEIN.		
GN	MPDZ.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57/BL6 X CBA F1; TISSUE=WHOLE BRAIN;		
RA	Simpson E.H., Suffolk R., Jackson I.J.;		
RT	"Identification and mapping of mouse Multiple PDZ domain protein,		
RT	Mpdz".		
RL	Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AJ131869; CAA10523.1; -.		
DR	HSSP; Q12959; 1PDR.		
DR	MGD; MGI:1343489; Mpdz.		
DR	InterPro; IPR001478; -.		
DR	Pfam; PF00595; PDZ; 13.		
DR	SMART; SM00228; PDZ; 1.		
SO	SEQUENCE 2055 AA; 218968 MW; E1A38EE634CC20EA CRC64;		

Query Match	84.8%	Score 4314.5	DB 11	Length 2055
Best Local Similarity	79.7%	Pred. No. 5e-252		
Matches	856	Conservative	71	Mismatches 74, Indels 73, Gaps 2
QY	2	LQWVSKESFERTININAKGSSSLGMYTVANSKDKGLMYVRSIIHGGALSIPDRGAIAGDGIUS	61	
Db	985	LQSMQEAEEFRRVTTAKGSSSLGMYTVANSKDKGLGYVRSIIHGGALSIRGRIRIIVGCIIS	1044	
QY	62	INEESTISVTNQARAMLRRHSLIGDITITYVPAHELEEFKSLGQSGRWALDIFSS	121	
Db	1045	INEESTISLTNQARAMLRRHSLIGDITITYVPAHELEEFKSPQOAGNIALDIFSS	1104	
QY	122	YTGROIPELPEREEGGESESLQNTAYSWMNOPRYVELRREPSKSLGISIVGRCMGSR	181	
Db	1105	YTGROIPELPEREEGGESESLQNTAYSSWSPRYVELRREPSKSLGISIVGRCMGSR	1164	
QY	183	SNGEVWRGIFIHVLEDSFAGKNGTLPEDRIIVE	215	
Db	1165	SNGEVWRGIFIHVLEDSFAGKNGTLPEDRIIVE	1224	
QY	216	-----APSGESEPEKAPLCS	231	
Db	1225	FMVQSTINRPRKSPLEPSHSLYPKYSFSSINPFADSLQLTQDQASQSESETEKALCN	1284	

Qy	232	VPPPPSAPFAMGSDHNOSSASXISODVYKDEDFEY5MKNIBERGTLNGELHMLTELK	291
Dd	1285	VPPSPSPVFSFMSGDCQAPSATVASEDEDEDFEY5MKNIDERGSLTGQHLVTELK	1344
Qy	292	HSGLGLSLAGKODSRMSASVTVIGIDPFGAAGKGRQIADDELLEINGQTLYGRSHQMS	351
Dd	1345	OSGIGLSLAGKODTRMSVTVIGIDPFGAAGRGRQIADDELLEINGQTLYGRSHQMS	1404
Qy	352	IICAPSKVKIIFIRNKDAVNQMAVCPGNAVEPLPSNSENLOKETEPTVTTSDAAYDLS	411
Dd	1405	IICAPSKVKIIFIRNADAVNOMAVCPGIADSPSSSTSDPNKEVEPCSTTSASAADLS	1464
Qy	412	SFKVWQHLELPKQOOGGIGIAISEEDPLSGYIISLTLREHVANTDGRLAKYDQIILAVDEI	471
Dd	1465	SLTVYQYOLELPKQDGGIGIAICEEDPTINGMIEISLTLREHGAARKDGRLKHGDHIIIAVDEV	1524
Qy	472	VGVPIEIEKIFSLTKATMYKLTIRHAPSOAVP5AAGAA5GEKKNSSOSLMPPOSGP	531
Dd	1525	VAGCPVEKEFISLTKTAATYKLTIVRAENPACPAVPSATYVSGERKDNQTPAV---AP	1581
Qy	532	EPESIRNTRSSTPAIPASDPATCPIIPGCETTIEISKGRTGJLSIVGSDTLGAPFI	591
Dd	1582	DLEPIPSRSRSTPAVPA5DPATCPIIPGCETTIEISKQGTGJLSIVGSDTLGATII	1641
Qy	592	HEVVEEGAACKDGLMAGDDQILEVNGIDLRKANHDAIVLNOTPORVRLTLXREDAEYK	651
Dd	1642	HEVVEEGAACKDGLMAGDDQILEVNGIDLRKANHDAIVLNOTPORVRLTLXREDAEYK	1701
Qy	652	EEBVCDDTLTIELOKPKPAGJLSIVGKRNTGVFNVDIYKGGIADPDGRILIOGDDILLVN	711
Dd	1702	EBDVCDDTLTIELOKPKPAGJLSIVGKRNTGVFNVDIYKGGIADPDGRILIOGDDILLVN	1761
Qy	712	GEDYRNASQEVVAALLKCSIGTYVTLVEYGRITAKGPHSERRPSOTSOVSEGSLSSTFFPLS	771
Dd	1762	GEDYRNHAQOEVAALLKCSIGATYVTLVEYGRYKAPAFHSERRPSOSSOVSESSLSSTFFPLS	1821
Qy	772	GSSTSESESSSKKNALASEIOGIRTYEMKKGPTDLSGISINGCVSPIGADVPITAMH	831
Dd	1822	GINSSESESSSKKNALASEIOGIRTYEIKKGADSLGISINGCVSPIGADVPITAMH	1881
Qy	832	PTGVAAQOTKLRVGDRIYVTLIGSTSEBMTQTQAVNLKLNKNA5GSIEMQVYAGDGVVWYGH	891
Dd	1882	PNGVAQAQOTKLRVGDRIYVTLIGSTSDCMHTQAVNLKLNKNA5GSIEMQVYAGDGVVWYGH	1941
Qy	892	HOEPASSLSFTGJLTSISITODDLGPPOCK5ITLERGPDGLGFTIVYG5GSPHGDPLIV	951
Dd	1942	QOELANPCLAFETGLTSSISIFPDGLGP5OSKITILDRGPDGLSPENIYVG5GSPHGDPLIV	2001
Qy	952	KTVPAKGAASDGLRLKRGDQIIANNGSLGCVYHEEVAVALIKRTKTYVTLML5	1005
Dd	2002	KTVPAKGAASDGLRLKRGDQIIANNGSLGCVYHEEVAVALIKRTKTYVTLML5	2055

RESULT	4			
008783				
ID	008783	PRELIMINARY;	PRF:	526 AA.
AC	008783			
DT	01-JUL-1997	(TREMBLrel. 04, Created)		
DT	01-JUL-1997	(TREMBLrel. 04, Last sequence update)		
DT	01-MAR-2001	(TREMBLrel. 16, Last annotation update)		
DE	90RF BINDING PROTEIN 1 (FRAGMENT).			
GN	9BP-1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus			
OX	NCBI_TaxID=10090;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=PANCREAS;			
RA	Lee S.S., Weiss R.S., Javier R.T.;			
RL	Submitted (Apr-1997) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AF000168; AAB57835.1; -			

DR HSSP: Q12959; 1PDR.
 DR Interpro: IPR001478; -
 DR Pfam: PF00595; PDZ; 4.
 DR SMART: SM00228; PDZ; 1.
 FT NON_TER 1
 SQ SEQUENCE 526 AA; 54948 MW; 3843C5E3AA210E8 CRC64;

Query Match 45.2%; Score 2297.5; DB 11; Length 526;
 Best Local Similarity 86.6%; Pred. No. 5.6e-11;
 Matches 457; Conservative 34; Mismatches 32; Indels 5; Gaps 2;

QY 480 FISLTKAKMTVKLTTHAENPDQAVPSAAGASGEKKNSQSILMVPQSGPESISINT 539
 DB 2 FISLTKAKMTVKLTTHAENPDQAVPSAAGASGEKKNSQSILMVPQSGPESISINT 58
 QY 540 SRSSTPAIFADPATCPIIPCECTTIEISKRTGLSLVSGSDTLGAFIIEYEEGA 599
 DB 59 SRSSTPAIFADPATCPIIPCECTTIEISKRTGLSLVSGSDTLGAFIIEYEEGA 118
 QY 600 ACKGRLMAGDQILEVNGIDLRKATHEAIVNLROTPORVLTLYRDEAPYKEEVCDTL 659
 DB 119 ACKGRLMAGDQILEVNGIDLRKATHEAIVNLROTPORVLTLYRDEAPYKEEVCDTL 178
 QY 660 TIE--LQKPKGGLSLVGRNDTGVFSDIVKGIADPDGRLIOGDILLVNGEDVRN 717
 DB 179 TIEHQLGRPKGLSLVGRNDTGVFSDIVKGIADPDGRLIOGDILLVNGEDVRN 238
 QY 718 ASQAVALLKCSLGYTLVEGRKAGPHERPSQTSQVSESSLSFTPLSGSSISE 777
 DB 239 ATQAVALLKCSLGYTLVEGRKAGPHERPSQTSQVSESSLSFTPLSGSSISE 298
 QY 778 SLESSKKNALASEIQLRTVEKKGPTDSIGISIAGVSGPLGDVPFIAMHPPTGVA 837
 DB 299 SLESSKKNALASEIQLRTVEKKGPTDSIGISIAGVSGPLGDVPFIAMHPPTGVA 358
 QY 838 QTKLRVGRDRIYTCISGTEGHTHOAVNLKKNASGSIEMOVYAGDVSVTGHQEPAS 897
 DB 359 QTKLRVGRDRIYTCISGTEGHTHOAVNLKKNASGSIEMOVYAGDVSVTGHQEPAS 418
 QY 898 SLSFTGTSITIFODDGPQCKSITLERGPDGFSIVGSGSPHGDLYIKTVFAK 957
 DB 419 PCLMFTGLTSSIFPDGPPQSKITLDRGPDGFSIVGSGSPHGDLYIKTVFAK 478
 QY 958 GAASEDGRKRDQIIAVNGQSLGVTHEEAVAILKRTKGTVTLMVLS 1005
 DB 479 GAASEDGRKRDQIIAVNGQSLGVTHEEAVAILKRTKGTVTLMVLS 526

RESULT 5
 043798 PRELIMINARY; PRT; 453 AA.
 AC 043798;
 DT 01-JUN-1998 (Tremblrel. 06, Created)
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE MULTI PDZ DOMAIN PROTEIN 1 (FRAGMENT).
 GN MUPPL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN.
 RX MEDLINE=98196865; PubMed=9537516;
 RA Ulmer C., Schumack K., Flege A., Lubbert H.;
 RT "Cloning and characterization of MUPPL, a novel PDZ domain protein.";
 RL FEBS Lett. 424:63-68(1998).
 DR EMBL: AJ001319; CAAD04680.1; -
 DR HSSP: Q12959; 1PDR.
 DR Interpro: IPR001478; -
 DR Pfam: PF00595; PDZ; 4.
 DR PIRam: PF00595; PDZ; 4.

DR SMART: SM00228; PDZ; 1.
 FT NON_TER 1
 SQ SEQUENCE 453 AA; 47423 MW; DA7915AD493C5D8 CRC64;

Query Match 44.3%; Score 2252; DB 4; Length 453;
 Best Local Similarity 98.2%; Pred. No. 2.5e-128;
 Matches 445; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 553 ATCPITPCCTTIEISGRGRLGLSLVSGSDTLGAFIIEYEEGAACKGRIMADQI 612
 DB 1 ATCPITPCCTTIEISGRGRLGLSLVSGSDTLGAFIIEYEEGAACKGRIMADQI 60
 QY 613 LEVNGIDLRKATHEAIVNLROTPORVLTLYRDEAPYKEEVCDTLYIELKPKGGLG 672
 DB 61 LEVNGIDLRKATHEAIVNLROTPORVLTLYRDEAPYKEEVCDTLYIELKPKGGLG 120
 QY 673 LSIYGRNDTGVFSDIVKGIADPDGRLIOGDILLVNGEDVRNASEAVALLKCSLG 732
 DB 121 LSIYGRNDTGVFSDIVKGIADPDGRLIOGDILLVNGEDVRNASEAVALLKCSLG 180
 QY 733 TVTLEVGRIKAGPHERPSQTSQVSESSLSFTPLSGSSISESSKKNALASEI 792
 DB 181 TVTLEVGRIKAGPHERPSQTSQVSESSLSFTPLSGSSISESSKKNALASEI 240
 QY 793 QGLRTVEKKKPTDSIGISIAGVSGPLGDVPFIAMHPPTGVAQTKLRVGRDRIYTC 852
 DB 241 QGLRTVEKKKPTDSIGISIAGVSGPLGDVPFIAMHPPTGVAQTKLRVGRDRIYTC 300
 QY 853 GSTEGMHTHOAVNLKKNASGSIEMOVYAGDVSVTGHQEPASSLSFTGLTSSIFQ 912
 DB 301 GSTEGMHTHOAVNLKKNASGSIEMOVYAGDVSVTGHQEPASSLSFTGLTSSIFQ 360
 QY 913 DDLGPPCKSTTLERGPDLGFSIVGSGSPHGDLYIKTVFAKGAASEDGRKRDQI 972
 DB 361 DDLGPPCKSTTLERGPDLGFSIVGSGSPHGDLYIKTVFAKGAASEDGRKRDQI 420
 QY 973 IAVNGQSLGVTHEEAVAILKRTKGTVTLMVLS 1005
 DB 421 IAVNGQSLGVTHEEAVAILKRTKGTVTLMVLS 453

RESULT 6
 09H3N9 PRELIMINARY; PRT; 1134 AA.
 AC 09H3N9;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE PDZ DOMAIN PROTEIN 3' VARIANT 4.
 GN HINADL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Kimura A., Okubo K., Mukai T.;
 RA "Cloning, characterization and chromosomal mapping of the two novel
 RT heart specific genes.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB044807; BAB19683.1; -
 SQ SEQUENCE 1134 AA; 125244 MW; FB6C3DA92D7CEFB6 CRC64;

Query Match 29.4%; Score 1496.5; DB 4; Length 1134;
 Best Local Similarity 44.3%; Pred. No. 5.3e-82;
 Matches 360; Conservative 99; Mismatches 177; Indels 177; Gaps 21;

QY 3 ONVSKESPER-----TINAKGSSSLGTMVSNKKGICLMIYSHHGALISRDGRIAGDC 58
 DB 402 ENVKKNFVMSLPSVSTSGNSQGR--FDDLENLSLAKTSLDGLKIPND----- 451

```

QY 59 ILINESTISVTNQAARMLRHSILGPDKITV-VPAREHLEEFKISIG--QSGRYMAL 116
DB 452 -----VQGPSLLIDLPVAAQREQEDLPYQOARVRISK 486
QY 117 DIFSSYTG-----RDIPELPEREGEGESELQNTATASNNOAPRYALMEEPSKSLGI 169
DB 487 --ASAVTGMSSRYATDICELEPEREGEGEET-----PNFSHNGPPRIYVIFREPNVSLGI 540
QY 170 STVGGRGMSRLSNGEVMGIFIKHYLEDSPAGKNGTLPGRDRIYEA-----PSQSESEP 224
DB 541 STVGQGVYIKRLKNGELGIFIKYLEDSPAGKNTALTKGILLEVSGVDLQNMASHSEA 600
QY 225 EKA-----PLCSVPPPPSAFAENG-----SDHTQSSASK----- 254
DB 601 VAIKNAGPVPVFIYOSTSTPVPVLPVHNKANKITGONODOTKEKKEKROGTAPPMKL 660
QY 255 -----ISODVDK-EDEFYSKMKINERYGTGLGELHMLEKSGLSGLSLAGNKRSR 307
DB 661 PPYKALTDSDSENEEDAFDTQKINQRYADLPGLHIIIELEKDKNGLSLAGNKRSR 720
QY 308 MSVFYIGIDPNGAGKDGRLQIADLELTNGOILYGRSHQNASITKCAPSKVYKIIETRN 367
DB 721 MSIFVYGINPEGPAAADGEMHIGDELLEINNOILYGRSHQNASAIKTAPSKVYKIVFIRN 780
QY 368 KQAVNOMAVCPGNAVEPLSENENLQNKETEPVYTSDAAVDLSPFKNVQHLLEPQDGG 427
DB 781 EAVNOMAVTP-----FPVSSSSP-----SST-----EDQSG 807
QY 428 LGAIASEEDTLGVLIKSLTEHGVATDGRKLVGDOILAVDEIVYGYPIEKFISILKTA 487
DB 808 TERPISSE-----DLSLEVGIKQLPESESFKLA-----VSGMKQO 842
QY 488 KMTVKLTTHAENPDQAVPSAAGASGEEKNSQSILMVPQSSPEPESIRNTSRSTPAI 547
DB 843 KPTKXSF-----SSQEIPLAPASS-----YHSTADFTGYGFGQAPLAV----- 882
QY 548 FASDPATCPTIIGCETTIEISKRGGLSLSTYCGSSTLLGAFTHIHYEREGAACKRGRLM 607
DB 883 ---DPTACTIVPGQEMIILISKRSGLSLSTYCGKDTPLNATVILHYEYEEGAARGRGLM 939
QY 608 AGDQILEVNGIDLKATKATHEAIVNLROTPOVRLTYRDEAPYKEEVECDTLTIELQKRP 667
DB 940 AGDQILEVNGVLDRLNSHHEAITALRQTPQKRLVYRDEAHYRDEENLEIFPVDLQKRA 999
QY 668 GGLGLSLSTYCGKRGNDGVFVSDIYKGIADPDGRLIOGDOILLVNGEDVNASQEAVALLL 727
DB 1000 GGLGLSLSTYCGKRGNSGVFSDIYKGAADLDGRLIOGDOILLVNGEDMKNASQETVATILL 1059
QY 728 KCSLGTVTLLEVGRIRKAGPFRHSRRPSQTSQVSE 760
DB 1060 KCAQGLVQLEIGRLRAGSWSATSQNSQSAE 1092

```

```

RESULT 7
ID 070471 PRELIMINARY: PRT: 612 AA.
AC 070471:
DB 01-AUG-1998 (TREMBLrel. 07, Created)
DB 01-AUG-1998 (TREMBLrel. 07, last sequence update)
DB 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DB CHANNEL-INTERACTING PDZ DOMAIN PROTEIN.
DE CIPP.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE-CEREBELLUM, BRAIN;
RA Kuznecher C., Mermelstein P.G., Holden W.T., Surmeier D.J.;
RL Mol. Cell. Neurosci. 0:0-0(1998).
EMBL, AF060539; AAC0148.1; -.

```

```

DR HSSP: 012959; 1PDR.
DR MGD; MGI:1277960; CAPP.
DR InterPro; IPR001478; -.
DR Pfam; PF00595; PDZ; 4.
DR SMART; SM00228; PDZ; 1.
SQ SEQUENCE 612 AA; 65420 MW; 145DE1769F54BE56 CRC64;

```

```

Query Match 28.9%; Score 1469.5; DB 11; Length 612;
Best Local Similarity 49.5%; Pred. No. 8.5e-81;
Matches 335; Conservative 85; Mismatches 158; Indels 99; Gaps 12;

```

```

QY 270 KNIRERYGTUTGELHMLEKSGLSGLSLAGNKRSRMSVYIGIDPNGAGKRGRLQI 329
DB 8 EKTIRQYADLPBELHIIIELEKDKNGLSLAGNKRSRMSITVGINPEGPAAADGRMTI 67
QY 330 ADELLEINQIILYGRSHQNASIIRKCAPSKVYKIIETRNDAVNOMAVCPGNAVEPLPSNS 389
DB 68 GDELEINNOILYGRSHQNASAIKTAPRYKLVFIRNEDAVSQMAVAP----- 116
QY 390 ENLQNKETEPVYTSDAAVDLSPFKNVQHLLEP-KDQGLGAIASEEDTLGVIYIKSLTE 448
DB 117 -----FPELSHPSPVEDLGCTELVSSSE--SSVDAKHLPE 151
QY 449 HGVAAATDGRKLVGDOILAVDEIVYGYPIEKFISILKTAQMVKLTTHAENPDQAVPSA 508
DB 152 PSS-----KEDLSQVYVDNMVAEQKE-----SESPDSACQIK 187
QY 509 AGASGEEKNSQSILMVPQSSPEPESIRNTSRSS-----TPAIRSPADPCTPIIP 559
DB 188 QQTYSQVSSSQD--SPSPAPLQOSAHADVTGSGNFQAPLPVDPAPLSTVPACPIYP 245
QY 560 GCETTIEISKRGGLSLSTYCGSDTLGAFTHIHYEREGAACKDGRMLAGDOILEVNGID 619
DB 246 GQEMIIEISKRSGLSLSTYCGKDPPLDAIVHYEYEEGAARADGRLMAGDOILEVNGVD 305
QY 620 LKATKATHEAIVNLROTPOVRLTYRDEAPYKEEVECDTLTIELQKRPKGLSLSTYGR 679
DB 306 LRSSSHEAIVNLROTPOVRLTYRDEAPYRDEAYRDEELEFLVLDLQKTRGGLSLSTYGR 365
QY 680 NDTGVFVSDIYKGIADPDGRLIOGDOILLVNGEDVNASQEAVALLLCSLGTVTLLEV 739
DB 366 SSGSVFISDLYVGAADLDGRLIRGDOILLVNGEDMKNASQETVATILKCYGVQVLEIG 425
QY 740 RIKAGPFRHSRRPSQTSQVSEGLSFTPLPGSSTSELES-----SSKNALASEIQ 793
DB 426 RIRAGSMAASRKTSONSQDQHSANSSCRP-SFAPVYTSQVLYGTRKSSDPPOKCTE-E 483
QY 794 GLRTVEMKGPVDSIGSLTAGVSPLDGVPPIFIAMHPTGYAAGTOKLRVGDRIYTCG 853
DB 484 EPTVEITIELSDALGSLTAGSGSPLDGPIPIFIAMIQANVGAARKOKLKVGDRIYSLNG 543
QY 854 TSTEGTHTQAVNLKLNAGSIEMQVYAGDVSVYTGHHQEPASSSTFTGLSTISIFOD 913
DB 544 QPLDGLSHTDVAVNLKLNAGGRILLYVADPTNLSALATOLEINSAGS----- 589
QY 914 DLGPPQCKSITLERGPD 930
DB 590 QLGSP-----TADRHP 601

```

```

RESULT 8
ID 060833 PRELIMINARY: PRT: 1552 AA.
AC 060833:
DB 01-AUG-1998 (TREMBLrel. 07, Created)
DB 01-AUG-1998 (TREMBLrel. 07, last sequence update)
DB 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DB INADL, C-TERM VARIANT2.
DE INADL.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

```



```

Db 1349 TEPISEE-----DGSLEVGIKQLPESSEFKLA-----VSOMKQ 1383
Oy 488 KMTVKLTTHAENPDQAVPSAAGAASGEKKNSSOSLWVPOGSPPEESIRNTSRSTPAI 547
Db 1384 KYPTKVSF-----SSQEIPLAPASS-----YHSTDAFTGYGFGQAPLAV-----1423
Oy 548 FASDPATPIIPGCTTTEISKRGTLGSLTVGSGDITL-----586
Db 1424 ---DPATCPPIVPGQEMITTEISKRGSLTVGSGKDTPLFWLGSPPRAMSOHLVRAFMILH 1480
Oy 587 -----GAFIIHEVEEGAACADKGRMLAGDQILEVNGIDLRKATDEAIVLRQTPQ 637
Db 1481 HPVTEVEGONALVIHVEVEEGAARDGRMLAGDQILEVNGVDLRNRSHEALTALRQTPQ 1540
Oy 638 RVRLLTVRDEAPYKEEYCDTLTIELQKPKGKGLSTVIGKR 679
Db 1541 KYRLVYRDEAHYRDEENLEIFPVLDQKKAGKGLSTVIGKR 1582

RESULT 10
ID 015249 PRELIMINARY; PRT: 1524 AA.
AC 015249;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PDZ DOMAIN PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97423468; PubMed=9280290;
RA Philipp S., Flockerzi V.;
RT "Molecular characterization of a novel human PDZ domain protein with
RT homology to INAD from Drosophila melanogaster.";
RL FEMS Lett. 413:243-248(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96224170; PubMed=8617505;
RA Lennon G., Auffray C., Polymeropoulos M., Soares M.B.;
RT "The I.M.A.G.E. Consortium: an integrated molecular analysis of
RT genomes and their expression.";
RL Genomics 33:151-152(1996).
DR EMBL; AJ001306; CAA04666.1; .
DR HSSP; P31016; 1BFE.
DR InterPro; IPR001478; .
DR Pfam; PF00595; PDZ; 8.
DR SMART; SM00228; PDZ; 1.
SO SEQUENCE 1524 AA; 167323 MW; E3BBA7EC856A954D CRC64;

Query Match 20.9%; Score 1063.5; DB 4; Length 1524;
Best Local Similarity 38.4%; Pred. No. 1.3e-55;
Matches 281; Conservative 87; Mismatches 155; Indels 205; Gaps 22;

Oy 3 QNVSKESEER---TINIAKNSLSLGMTVSAKNGDLGMLVRSIIHGALSRDGRITAGDC 58
Db 943 EKVMEKENVMESLPSVPSTEGNSQGR--FDLEMLNLSIAKSLDLMIPND-----992
Oy 59 ILSINEESTISTYNAQARMLRRLSLGPDIKITY-VRAEHLERKISLGG-QSGRWAL 116
Db 993 -----VQGSLLIDLPVVAQREQDLPYQHQAARVLSK 1027
Oy 117 DISSSYTG-----RDIELPERECEGESEELQNTAVSNMNPRLVRLMREPSKSLCI 169
Db 1028 ---ASAVTGMLSRVAITDCELPERECEGESE---PNFSHMGPRIVAEIFRPNVSLGI 1081
Oy 170 STVGRGMSGRSLNSGEVNMKGIFIKVHLEDSPAKNGKTLKPGDIRVEA-----PSGSESEP 224
Db 1082 STVGQATYKRLKNGEELKGIKIVOLEDSPAKNTALKTGDKILEVSGVDLQNSHSPA 1141

```

```

Oy 225 EKA-----PLCSVPPPPSAFAEMG-----SDHTOSSAK-----254
Db 1142 VEAIRNAGNPVVEIVQSISSTPRVIPNVNANKITGQNONDQTKKRRQGTAPPPKL 1201
Oy 255 -----ISODVDR-EDEFGYSWKINRERYGLTGLHIELEKSGSLGSLAGKDKSR 307
Db 1202 PPPYKALDDSDSENEEDAFIDQKIRQRYADLPGLHLIIELEKNGSLGSLAGKDKSR 1261
Oy 308 MSVFLVIGDIPNCAAGKDRLOJADELLEINQOILYGRSHQNASIIKCAPSKVLIIFIRN 367
Db 1262 MSIFVGVINPEGPAAADGRMHIGDELLEINQOILYGRSHQNASIIKCAPSKVLIIFIRN 1321
Oy 368 KDVAVQMAVCQNAVEPLPSNSENLQNKETEPTVTSDAADLSSEFKVQVHLEPKDDG 427
Db 1322 EDVAVQMAVTP---FPPSSSP-----SSI-----EDSG 1348
Oy 428 LGIAISEDTLSGVITIKSLTEHGVAATDGRKLVGQDILAYDDEIVGVPIEKFISLTKTA 487
Db 1349 TEPISEE-----DGSLEVGIKQLPESSEFKLA-----VSOMKQ 1383
Oy 488 KMTVKLTTHAENPDQAVPSAAGAASGEKKNSSOSLWVPOGSPPEESIRNTSRSTPAI 547
Db 1384 KYPTKVSF-----SSQEIPLAPASS-----YHSTDAFTGYGFGQAPLAV-----1423
Oy 548 FASDPATPIIPGCTTTEISKRGTLGSLTVGSGDITLGAFTIHEVEEGAACADKGRML 607
Db 1424 ---DPATCPPIVPGQEMITTEISKRGSLTVGSGKDTPL-----1459
Oy 608 AGDQILEVNGIDLRKATDEAIVLRQTPORVRLTVRDEAPYKEEYCDTLTIELQKRP 667
Db 1460 -----VNGVDLRNRSHEALTALRQTPQKRYLVYRDEAHYRDEENLEIFPVLDQKKA 1512
Oy 668 GKGGLSTVIGKR 679
Db 1513 GRGLSTVIGKR 1524

RESULT 11
ID 009515 PRELIMINARY; PRT: 2208 AA.
AC 009515;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE HYPOTHEtical 240.1 KDA PROTEIN C52A11.4 IN CHROMOSOME II.
GN C52A11.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodertinae; Caenorhabditis.
NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL NZ;
RA Coles L., Sulston J.;
RT Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z46792; CAA86769.2; JOINED.
DR EMBL; Z46795; CAA86769.2; JOINED.
DR EMBL; Z46792; CAA86789.2; JOINED.
DR HSSP; Q12959; 1PBR.
DR InterPro; IPR001478; .
DR InterPro; IPR002173; .
DR Pfam; PF00595; PDZ; 15.
DR PROSITE; PS00584; PKB_KINASES_2; UNKNOWN_1.
SO SEQUENCE 2208 AA; 240120 MW; B6894F8D603E79FB CRC64;

Query Match 16.0%; Score 814; DB 5; Length 2208;
Best Local Similarity 20.98%; Pred. No. 2.9e-40;
Matches 307; Conservative 137; Mismatches 259; Indels 768; Gaps 33;

Oy 11 ERTINIAKNSLSLGMTVSAKNGD-LG-LGMIVRSIIHGALSRDGRITAGDCILSINEESTI 68

```


FT MUTAGEN 188 188 Y->F: NO EFFECT ON BINDING TO NUMB
 FT VANSPLIC 1 131 PROTEIN
 FT DLIHCICQALLDPLDTCGHTGTYCLTCLNLEVEDCPVD
 FT RRPVLOHCKRSSILVNLKLNKLTATCTCTCTEGLORCD
 FT LOHHEOTS -> MKALLLVLPWLSPPANVIDVGLHRLYS
 FT EL (IN ISOFORM LNX-B)
 FT SEQUENCE 728 AA: 80156 MW: E2914BD364C0CEC4 CRC64;

Query Match 11.6%; Score 589; DB 11; Length 728;
 Best Local Similarity 29.1%; Pred. No. 2e-27;
 Matches 188; Conservative 115; Mismatches 239; Indels 106; Gaps 23;

QY 411 SSKFVQHLLEPKD-----QGGL--GIAISEEDTISGVITKSLTEHGAATDGRKLKVGDO 463
 DB 130 TSCKASHGHLTKDRKRSQDCPDGCASLMTLIS-----PEVSAAT----- 173
 QY 464 ILAVDEIVGYPIEKELSLTKAKMTVLTJHAENPOQAVP---SAGAASGEKKNS 520
 DB 174 ISLMTDEPGLNP--AYSSVEDGEDEVANSSDGSNKRTRARPERSTKRSFKINRA 231
 QY 521 QS-LMWPOSGSPEPSI-----NTSKRSPTAFASDPATCPPIPGCE-TTIEISKGRGL 574
 DB 232 LSLARKTKSGYVAVNHVDQGRDNSEHTVEYF---PRLFHLIPDELTSIKINRADPSE 288
 QY 575 GLSI--VGSOTLLGAFITHEVEBACDKDRIMAGDQILVNGDILKKAHDEAINLV 632
 DB 289 SLSIRLVGSESLPVLHIIITHTYRGVIAARDRLRGLDILVNGDISNVPHNVAVRL 348
 QY 633 RQTPORVRLTYRDE-----APYKEEVECTLTIELQK-KPKGLGSLVIGKRRD 682
 DB 349 RQPCOVLRLTYRDEKFRSNAHNPDSYGRPDSDHVLTKNSSPEQGLITLVRVDEP 408
 QY 683 GVFSVDIVKGIADPDGRLLIQGQDILLVNGEDVRNASSQGAVALAKSLGVTLEVGRIT 742
 DB 409 GVFIENLVNGVADRHRQGLENDRLVLAINGHDLRFQSPSAHLIQASBRVHLVYSR-- 466
 QY 743 AGRFHSERPPS-----QTQVSEGLSSTFTPLSGSSTESLESSKKAALSEIOGLT 797
 DB 467 -----QVRQSPDIFQEACWISNGOOS-----PQGER-----NTASKRPATCHE---KV 508
 QY 798 VEMKKGPTSLGISIAGVSPGLDVPFIAMNHPTGVAQOTOKLEGRITYICTSTRE 857
 DB 509 VSMKMDPSLSLGMTVGGGASHREMDLPITYIVSEPGVTSRGRITGIGILLANCIELT 568
 QY 858 GMTHTQAVNLLKNASGISIMQVAVAGDVSVTGHQEPASSLSLFTGLTSTIFODDLCP 917
 DB 569 EVSRTAEVAVLKSAFSSVVLKAL---EVK-----EQEAQEDCSPALDS---NHNVT 615
 QY 918 P-----QCKSTLERGPDG-LGFSYVGGVSGPRGDIPIYVKYFAAG 958
 DB 616 PEDWSPSWMMLELPOLYLCNCKDVLIRNTASGLFCYGVGEYSGNPFPIKTSIVEGT 675
 QY 959 AASEGRLKRGDQILAVNGQSLEGVTHEBAVALIKRTKGTVLMLVS 1005
 DB 676 PAVNGCRIRCGDILLAVNGRSTSGMTHACLAKMLKELKRITLTIAS 722

RESULT 13
 Q23823 PRELIMINARY: PRT: 665 AA.

AC 023823: 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE INAD PROTEIN.
 GN INAD.
 OS Calliphora vicina (Blue blowfly) (Calliphora erythrocephala).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Oestroidea; Calliphoridae; Calliphora.
 OC NCBI_TaxID=7373;

RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=CHALKY; TISSUE=RETINA;
 RX MEDLINE=96216082; PubMed=8662634;
 RA Huber A., Sander P., Paulsen R.;
 RT "Phosphorylation of the inad gene product, a photoreceptor membrane
 FT protein required for recovery of visual excitation."
 RL J. Biol. Chem. 271:11710-11717(1996).
 CC -I- FUNCTION: MAY BE INVOLVED IN CONTROL OF THE LIGHT RESPONSE. A RISE
 CC IN INTRACELLULAR CALCIUM LEVELS UPON VISUAL EXCITATION MAY
 CC INITIATE PHOSPHORYLATION OF THE INAD PROTEIN BY EYE-PKC.
 CC PHOSPHORYLATED INAD MAY IN TURN ACT ON ANOTHER PROTEIN SUCH AS TRP
 CC OR NORPA WHICH BOTH CO-PRECIPIRATE WITH INAD.
 CC -I- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED.
 CC -I- TISSUE SPECIFICITY: EXPRESSED IN RHABDOMERAL PHOTORECEPTORS.
 CC -I- SIMILARITY: TO DROSOPHILA INAD.
 DR EMBL: Z69889; CAA93758.1;
 DR HSSP: P31016; 1BFE.
 DR InterPro: IPR001478;
 DR Pfam: PF00595; PDZ; 5.
 DR SMART: SM00228; PDZ; 1.
 KW Repeat, phosphorylation; glycoprotein.
 FT DOMAIN 41 420 2 x APPROXIMATE REPEATS.
 FT REPEAT 41 82 1.
 FT REPEAT 381 420 2.
 FT CARBOHYD 5 5 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT MOD_RES 15 15 BY PKC (POTENTIAL).
 FT MOD_RES 168 168 BY PKC (POTENTIAL).
 FT MOD_RES 185 185 BY PKC (POTENTIAL).
 FT MOD_RES 320 320 BY PKC (POTENTIAL).
 FT MOD_RES 321 321 BY PKC (POTENTIAL).
 FT MOD_RES 437 437 BY PKC (POTENTIAL).
 FT MOD_RES 595 595 BY PKC (POTENTIAL).
 FT MOD_RES 658 658 BY PKC (POTENTIAL).
 SQ SEQUENCE 665 AA: 73349 MW: C453192E18202BF7 CRC64;

Query Match 11.1%; Score 564.5; DB 5; Length 665;
 Best Local Similarity 26.8%; Pred. No. 5.3e-26;
 Matches 190; Conservative 109; Mismatches 229; Indels 181; Gaps 27;

QY 153 QPRRVLEMRPSKSLGISIVGRGMGSRLSNGEVMKGIPIKHYLEDSPAGKNGLKPGDR 212
 DB 10 QVOSTYLDKTKGKSGSLVRGBE-----RDSNSKGIPIKGIYVDPSPGHLGKIKVGR 64
 QY 213 IYEARS---QSESEPE----- 225
 DB 65 LTLNGKDVDRDTEPEVINLIKQAGSKIDLEQTYGSEQSNKMGMEIKENGESQNRNMEN 124
 QY 226 ---KAPLCVSP-----PPPSAFAEMGSDHTOSSAS----- 253
 DB 125 QDSINQPTKQPAIKQOSMKQAPGRPRVNNAMKNSNTTSSKSDQDLDDDEDTEDTDM 184
 QY 254 -----KISQVDK--EDEFYSMKNIRERYGTGLGELHMIEL 288
 DB 185 TGRIRTAGVEIDRASAGCNKLNKMEKDRDKETDEDFGTAKINKRYNTMR-DLKKEI 243
 QY 289 EK-GHSGGLSLAGNDRKRSMSFYIVGIDPNG-AAGKDRLOIADLELFINQIILYGRSH 346
 DB 244 VPTNTALTLAAGSHDRKMGCFVAGVNTSGPLASVD--IKSGEILTEVNGTIVLKNRCH 301
 QY 347 ONASSITIKCAPSKVKIIFRNKDAVNOMAVCPGNAVEPLPSNSENLONKETPTVTTSDA 406
 DB 302 LNASVTFKNIDG-RVLITSRRKRPND---EGMSVKPI-----KKPPEI--DDT 345
 QY 407 AVDLSSFKNVQHLLEPKDQGLGIAL--SEBDTISGVITKSLTEHGAATDGRKLKVGDO 463
 DB 346 KFLFEQYAKARASVOY-KKRGFLGIMVIYGVKHYEVCNGIFISDLRESNMLAG-LKVGDM 403

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Sidenklamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Morley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zheng G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-COMPOUND EYE.
 RA Lindsley D.L., Zimm G.G.;
 RL "The genome of *Drosophila melanogaster*";
 RT Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-COMPOUND EYE.
 RA MEDLINE-9512723; PubMed-7826638;
 RA Shieh B.H., Niemeyer B.;
 RT "A novel protein encoded by the *Inad* gene regulates recovery of visual
 RL Neuron 14:201-210(1995).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE-COMPOUND EYE.
 RA Shieh B.;
 RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE-COMPOUND EYE.
 RA Shieh B.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE003458; AAF46915.1; -
 DR EMBL: U15803; AAC36490.1; -
 DR HSSP: P31016; 1BFE
 DR FlyBase; FBgn0001263; *Inad*.
 DR InterPro; IPR001478; -
 DR Pfam; PF00595; PDZ; 5.
 DR SMART; SM00228; PDZ; 1.
 SO SEQUENCE 674 AA; 74332 MW; DAC24091D99EAF73 CRC64.

Query Match 10.8%; Score 551.5; DB 5; Length 674;
 Best Local Similarity 26.3%; Pred. No. 3.3e-25;
 Matches 197; Conservative 133; Mismatches 285; Indels 135; Gaps 30:

OY 19 GNSSLGMY-----SANKDLGMYRSIIHGALSRDRIAGDCLISNEESTISVT 71
 DB 24 GKRSFGICIVRGVDSPTKTGTGIFIGIVDSPAHLGRLKVDRLISLNGK--DVR 80
 OY 72 NAQARAMLRRHSLIGPDIKITVPAPHLLEEFISLGQSGRYMALDIFSSTGRDIPLEP 131
 DB 81 NSTEDAVI---DLIKE-----ADFKIEL-----ETQTFD----- 106
 OY 132 EREEGESESLOMTAYSNMNPARRVELMRPEPSKSLGISIVGRGMSRLSNGEVMRGIF 191
 DB 107 -----KSDGQAKSDPRNGNGTMAKKNKFNQEDTNNNAS--GGCGMGCGGCGGCMAGM- 158
 OY 192 IKHVLDSPPAKNGTLPKGDRIVEAPSOSESEPEKAPLCSVPPPPSAFAEMGSDHTOSS 251
 DB 159 ---NRQSGMOKRNTTFTASMKOKHSNVADEDEDETRDMTG-----RIRTEAGYEIDRAS 209

OY 252 A-----SKISODVDK--EDERFGYSWKNI RERYGLTGLHMELEKGS-GGLSLAGNK 303
 DB 210 AGNCKLNQOEKDRKQEDDEFEGTYMAKINKRYNMKK-DLRRIEVRDASKPLGLLACHK 268
 OY 304 DRSRMSVPIVIGIDPNGAAGKGRLOIADELLEINGOILYGRSHONASSIIICAPSRYII 363
 DB 269 DROKMACFVAGVDPRGALGSVD-IKPGDEIYEVNGNVLNKNCHLMAVASVFNKVDGD-KLV 326
 OY 364 FIRKDAVNOAIVCGNAVEPLPSNSENLOKKEPEPTYSDAVDSFFKNVGHLELPK 423
 DB 337 MITSRRKPNDEGMC---VKPI-----KKPPTASDETKFI-FDQFPKARTVQVRK 371
 OY 424 DOGGIGIAT---SEEDTLGVIKSLTEHGAATDGRKVGQDILAVDDEIVGYPIEKF 480
 DB 372 -EGFLGIVITYGKHAVESSGIFISDLREGSNAELAG-VKVGDMILAVNODVLTESNYDA 429
 OY 481 ISLKTAKMYK--LTHAENPDSQAVPSAAGAAGEKSSOSGLWPVQSGSPESIR 537
 DB 430 TGLTKRAGVVTMILLTSE---EAIKAEKAEKKEEKKKEEPQ----- 475
 OY 538 NTSRSTPAIRASDPATCPPIPGCETTIEISKRGGLSLVGSQD--TLGAFIIEHY 595
 DB 476 -----EPATAEIKPNKKILLELVEKRPNGVIVCGKNHVTTCVITH-VY 521
 OY 596 EGAACKDGRMLWAGQDILEVNG--IDLKRTATHEAIVNLRQTPOR-VRLTYRDEAPYKE 652
 DB 522 PEGOVAAPKRLKIPHCIDINGTPIHVGSMITLAKVHOLFHTTEKAVTLVFRADPELE 581
 OY 653 EEVCDTLTIELOKKPKGKGLSLVSGKRNQDGVFVSDIYKGGIADDPGRILIOGQDILLVNG 712
 DB 582 K-----FNVDLMKRKAGKEGLSL--SPNEIGCTIADLIOGQVPEIDSKLQRDITTKNG 634
 OY 713 EDVRNASEAVALKCKSLGTYTLEVGRIK 742
 DB 635 DALEGLPVOVCYALFKGANGKVSMEVTRPK 664

Search completed: July 12, 2001, 14:43:13
 Job time: 226 sec

Fri Jul 13 15:00:14 2001

us-09-502-698-2.rpt

Page 12

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2001, 14:40:58 ; Search time 23.36 Seconds
(without alignments)
866.668 Million cell updates/sec

Title: US-09-502-698-2
Perfect score: 5085
Sequence: 1 MGNVSKESFERTINIAKGN.....EEAVALKRKRTVTLWLVS 1005

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PTCUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/Backfilest1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	509	10.0	2485	4 US-09-290-640-46	Sequence 46, Appl
2	500.5	9.8	2466	3 US-09-080-855-12	Sequence 12, Appl
3	500.5	9.8	2466	5 PCT-US94-09943-2	Sequence 2, Appl
4	498	9.8	2465	2 US-08-596-281-3	Sequence 3, Appl
5	498	9.8	2465	3 US-09-100-804-3	Sequence 3, Appl
6	340	6.7	1112	3 US-09-045-632-2	Sequence 2, Appl
7	340	6.7	1112	3 US-09-045-632-3	Sequence 3, Appl
8	335	6.6	702	3 US-09-045-632-15	Sequence 15, Appl
9	335	6.6	1018	3 US-09-045-632-16	Sequence 16, Appl
10	335	6.6	1061	3 US-09-045-632-32	Sequence 32, Appl
11	334.5	6.6	610	1 US-08-410-804-1	Sequence 1, Appl
12	334.5	6.6	610	1 US-08-259-514-1	Sequence 1, Appl
13	334.5	6.6	610	2 US-08-858-311-1	Sequence 1, Appl
14	334.5	6.6	1050	3 US-09-045-632-50	Sequence 50, Appl
15	331.5	6.5	1050	3 US-09-045-632-49	Sequence 49, Appl
16	329.5	6.5	604	3 US-09-045-632-14	Sequence 14, Appl
17	327	6.4	918	3 US-09-045-632-21	Sequence 21, Appl
18	327	6.4	961	3 US-09-045-632-33	Sequence 33, Appl
19	321	6.3	602	3 US-09-045-632-20	Sequence 20, Appl
20	288.5	5.7	861	3 US-09-045-632-34	Sequence 34, Appl
21	287	5.6	818	3 US-09-045-632-25	Sequence 25, Appl
22	273	5.4	507	3 US-09-045-632-13	Sequence 13, Appl
23	271	5.3	504	3 US-09-045-632-19	Sequence 19, Appl
24	266.5	5.2	642	3 US-09-045-632-35	Sequence 35, Appl
25	265	5.2	599	3 US-09-045-632-28	Sequence 28, Appl
26	258.5	5.1	502	3 US-09-045-632-24	Sequence 24, Appl
27	252	5.0	374	4 US-09-091-405-2	Sequence 2, Appl

28	227	4.5	519	3 US-08-997-445D-2	Sequence 2, Appl
29	225.5	4.4	283	3 US-09-045-632-27	Sequence 27, Appl
30	219	4.3	498	3 US-09-045-632-30	Sequence 30, Appl
31	219	4.3	541	3 US-09-045-632-36	Sequence 36, Appl
32	217	4.3	284	3 US-09-045-632-12	Sequence 12, Appl
33	214.5	4.2	407	3 US-09-045-632-18	Sequence 18, Appl
34	209.5	4.1	1612	3 US-08-545-860D-48	Sequence 48, Appl
35	209.5	4.1	1612	5 PCT-US94-04496-48	Sequence 48, Appl
36	203	4.0	404	3 US-09-045-632-23	Sequence 23, Appl
37	200	3.9	79	3 US-09-100-804-27	Sequence 27, Appl
38	200	3.9	86	5 PCT-US94-04496-53	Sequence 53, Appl
39	200	3.9	86	5 PCT-US94-04496-53	Sequence 53, Appl
40	196.5	3.9	198	3 US-09-045-632-11	Sequence 11, Appl
41	190.5	3.7	1829	4 US-09-157-420-1	Sequence 1, Appl
42	178.5	3.5	182	3 US-09-045-632-29	Sequence 29, Appl
43	176.5	3.5	233	2 US-09-151-611-1	Sequence 1, Appl
44	173	3.4	505	1 US-08-123-161A-14	Sequence 14, Appl
45	173	3.4	505	1 US-08-483-278-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-290-640-46
Sequence 46, Application US/09290640
Patent No. 6204055
GENERAL INFORMATION:
APPLICANT: Dean, Nicholas M.
APPLICANT: Marcussen, Eric G.
TITLE OF INVENTION: Antisense Compound Modulation of Fas Mediated Signaling
FILE REFERENCE: ISPH-0351
CURRENT APPLICATION NUMBER: US/09/290, 640
CURRENT FILING DATE: 1999-04-12
NUMBER OF SEQ ID NOS: 85
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 46
LENGTH: 2485
TYPE: PRT
ORGANISM: Homo sapiens
US-09-290-640-46

Query Match	10.0%	Score 509;	DB 4;	Length 2485;
Best Local Similarity	22.8%	Pred. No. 1.1e-34;		
Matches 243;	Conservative 140;	Mismatches 371;	Indels 314;	Gaps 37;
OY	162	EPKSLIGISIVGRMGSRNLSNGEVMRGIFIKHYEDSPAGKNGTLKPGDRIVEAPSOE	221	
DB	990	EPPPTVAELVYKPSHQRSDAESLAGY---TKLNKSVASLNKSPERRKHESDSSI	1046	
OY	222	SEPERAPLCVPPPPPSAFEMGSDHTOSASAKISQDYDKEDFEGYSWKNIERYGTTLG	281	
DB	1047	EDPGAVYLGMT-----MHSSGSSQVPLKENDV-----LHKRMSIVSS	1086	
OY	282	---ELHMELEK-GHSGIGLSIAGKKDSR--MSFYIGIDPNGAAGDKRGQIDELLE	335	
DB	1087	PERETTLVNLKRDARYGGLFQIIGKKMRDLGIFISSVAGGADLDGCKLPDRILIS	1146	
OY	336	INGOILYRSHONASSIICAPSKVIFIRNKDVMONAVPCGAAPVLPNSNENLNK	395	
DB	1147	VNSVLEGVSHHAALEIIONAPEDVTIVISQPKKISKVPSPVH---LTKEMKNYKK	1202	
OY	366	ETEPYVTTSDAAVDSFRN-----VOHLELPK--DOGLT--GIAISEDTLSVIKIS	445	
DB	1203	SS---YMDSKIDSSKDHHSRGLTNHISNSPGGLREGSLSDSDSTESASLSQ	1258	
OY	446	LTEHVAATDGLTKKGDOLAVDDIIVGVPIEKISILKTRAKMVKLTJHAENDSQAV	505	
DB	1239	SOVNEFFAS---HLGDQTW---QESHGSPSPVIS-----KATEKETFDSDNSKTRK	1306	
OY	506	PSAAGAASGEKK-----NSQSILVPOSGSPESIRNTSRSTPAIFADPATYC	555	

```

Db 1307 PGIDVYDYSRGSNDDEATYSSSDHQPCKO---ESSSVNTSNKMFTESSSPK- 1362
QY 556 PIPGCEITIEISKRGGLGSIYGSPTLL--GAFIHEVEEGACCKDRMLAGDOIL 613
Db 1363 ---PGDIFEVLANNDNSLGSIVGVNTSVRHGIYKAVIPQGAESDRIRKGRVL 1419
QY 614 EVNGIDLKATKTHDAINLVROTQORVRLTYRDEAPYKEEV---C----- 656
Db 1420 AVNGVSLGATVHAKOAVETLRNTGQVNHLLLEKGSPTSKEHVPTPQCTLSDNAQGGP 1479
QY 657 -----DTLIELOKKPKGGLS-----IVGRNDGVVSDIVK 691
Db 1480 EKVKKTTQVKDYSVTEENTEVVKFN--SSGLGFSRREDNLLPEQINASIVYKLLP 1538
QY 692 GGIADPDGRLIGDOILLVNGEDVRNASSDEAVALLKCSLTGLTEVGRVAG----- 744
Db 1539 GQPAAESKIDVGVILKVNASLKSQGVYISALRGTAEVLLCRPPGVLPEDT 1598
QY 745 ---PFHSE-----RRSOTSQVSEGS-- 762
Db 1599 ALLPPLSPACVLPNSSKSSQPCVEOSTSDENEMSDKSKKCKSPSRDYSDDSGS 1658
QY 763 -----LSFTF-----PLSGSST----- 775
Db 1659 GEDDLVYAPANISNYSWSSALLHQTLSNVAQASHHEAPKSOEDTICTMYYPQKIPNK 1718
QY 776 -----SESLESSK-----KNALASEI 792
Db 1719 EFEDSNSPRLPDPMAFGOSYPOSESASSSMKXHHIHSEPTROENMTLAKNDLEHL 1778
QY 793 QGLR-----IVEMKKGPTDSLGISIAG--VGSPLGDVPFIAMHPTGVAAOTOKLRV 844
Db 1779 EDFEVELLITLTKSEKSLGFTVTKGNQIRIGYVHV-----IQDP--AKSDGRLRK 1830
QY 845 GDRIVTICGTBEHTHQAANLLKNASGSIEMOVVAGGVVYTGHHOE-PASSLSFT 903
Db 1831 GDRILKNDVDTVMTHTDAVNLRAASKTVRL-----YIGRVLELPRIPLM-- 1877
QY 904 GLTSTIFODDLGPPCKSITLERGPDGLGFSIVGYSPPHGLDPIYKTVFAKAASED 963
Db 1878 -----PHLPDITLTCNKELESGSGHDSLY--QVVISDINPRSVAAIE 1922
QY 964 GLRKGDOIIYVNGOSLEGVTHEEA-----VAIKRTKGTVTML 1002
Db 1923 GNLDLIVHYVNGVSTGCMTLEEVNRALDMKSLPVLKATRNDLPVY 1970

RESULT 2
US-09-080-855-12
: Sequence 12, Application US/09080855A
: Patent No. 6083721
: GENERAL INFORMATION:
: APPLICANT: Saras, Jan
: APPLICANT: Franz, Petra
: APPLICANT: Aspenstrm, Pontus
: APPLICANT: Hellman, Ulf
: APPLICANT: Genez, Leonel Jorge
: APPLICANT: Heidin, Carl-Henrik
: TITLE OF INVENTION: PARC. A GTPASE ACTIVATING PROTEIN WHICH INTERACTS WITH PTPPLI
: FILE REFERENCE: 10461/7030
: CURRENT APPLICATION NUMBER: US/09/080,855A
: EARLIER FILING DATE: 1998-05-18
: EARLIER APPLICATION NUMBER: 08/805,583
: NUMBER OF SEQ ID NOS: 39
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 12
: LENGTH: 2466
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-080-855-12

```

```

Query Match 9.8%; Score 500.5; Db 3; Length 2466;
Best Local Similarity 22.2%; Pred. No. 6.2e-34;
Matches 255; Conservative 146; Mismatches 396; Indels 351; Gaps 41;

QY 142 ELQATASNNQPRVVELMREPSLSGISYGGKMSRLSNG----- 184
Db 868 QLOMRAROSNODADIERASFRLNLQAESYRGFMGALISTGLASTYLNKLAVPLSV 927
QY 185 --EVMRGIFIKHV-----LEDSPAGKNGTL---KPGD--RIVAPQSESEPEKAPLCS 231
Db 928 QAEILKRLSCSELQYLOPLONSSKEKNDKASWEKPREMSKSYNDLSQASLYPIRKNAVY 987
QY 232 VPPPPSAFAM--GSDHTQSSAS-----KISODVAKED 263
Db 988 NMEPPQGVAVELVYKPSHQMSRSDAESLACVTKLNKSVASLNRSPERRKHEDSSIE 1047
QY 264 EFG--YSKNIRERYGTLTG--ELHMLEEK--GHSGLGLSLACNKRDR--MSVFYVGI 315
Db 1048 DPGQAVYLDVYHKRMSYVSPEREITLVNLKDKAKYGIGFQIIGEKMGRLDGLIFISV 1107
QY 316 DPNCAAGKDRLOIADLELLEINGOILYGRSHONASITIKCAPSKVKTIFIRNKDVAVNOA 375
Db 1108 APGGPADRHGCLKPGDRLLISVNSVSLGVSHHAAIEILDQNPEDVTYISQPKKISKVP 1167
QY 376 VCPGNAAVEPLPSNSENLOKTEPEVTTSDAVADLSEFKN-----YQHLELPK--DOGG 427
Db 1168 STPVH-----LTNEMKNMTMKSS-----YMDSADISSCKDHMSGTLRHISENFGSPSG 1219
QY 428 L--GIAISEEDTSLGVYIKSLTEHGVAAIDGRKLVGDOILAVDEIYVGVPIEFISLKL 485
Db 1220 LRESLSSODSRFESASISQSVNGVFPAS--HAGDQTV--OESOHGSPSPVIS-- 1269
QY 486 TAKTVTLTHAENPDSQAVPSAAGAASGEKK-----NSQSILAMPQSSPEPES 535
Db 1270 --KATEKETFDLSQOSTKPKGISDVYDYSRGSNDDEATYSSSDHQPCKO---ESSS 1324
QY 536 IRTSRSTPAIFASDPATPCPIIPCETIEISKRTGLSIVGSDTLL--GAFIHE 593
Db 1325 SVNTSNKMFTESSSPK---PGDIFEVLANNDNSLGSIVGVNTSVRHGIYVKA 1380
QY 594 VYEGCAACKORMLAGDOILLEVNGIDLRKATHDAINLVROTQORVRLTYRDEAPYKEE 653
Db 1381 VTPQGAESDGRIRHKGDRVLAVNGVSLGATVHAKOAVETLRNTGQVNHLLLEKGSPTSKE 1440
QY 654 EV-----C-----DTLIELOKKPKGGLS----- 674
Db 1441 HVPYTPQCTLSDNAQGGPPEKVKTTQVNDYSVTEENTEVVKFN--SSGLGFSRSRE 1499
QY 675 ---IVGRNDGVVSDIVVAGGIADPDGRLIGDOILLVNGEDVRNASSDEAVALLKCSL 731
Db 1500 DNLIPEOINASIVYKRLFGQPAESGKIDVGVILKVNASLKSQGEVIALRGTA 1559
QY 732 GTVTVLEVGRIKAG-----PFHSE----- 749
Db 1560 PEVFLLCRPPGVLPEDITALLTPLOSPOVLPNSSKSSQPCVBOSTSDENEMSDK 1619
QY 750 ---RRPSOTSQVSEGS-----LSFTF-----PL 770
Db 1620 SKKCKSPSRDYSDDSGSEDDLVYAPANISNYSWSSALLHQTLSNVAQASHHEAPK 1679
QY 771 SGSST-----SESLESSK----- 784
Db 1680 SOEDYTCTMYYPQKIPNKPEFEDSNSPRLPDPMAFGOSYPOSESASSSMKXHHIH 1739
QY 785 -----KNALASEIQGLR-----IVEMKKGPTDSLGISIAG--VGSPLGDVP 824
Db 1740 SEPTROENMTPLKNDLEHLEDEFEVELLITLTKSEKSLGFTVTKGNQIRIGYVHV- 1798
QY 825 IFIAMHPTGVAAOTOKLRVGRIVTICGTBEHTHQAANLLKNASGSIEMOVVAGGD 884
Db 1799 ---IQDP--AKSDGRLRKQDRILKYNVDVMTHTDAVNLRAASKTVRL----- 1844
QY 885 VSVYTGHHOE-PASSLSFTGLSTSIPODDLGPPOCKSITLERGPDGLGFSIVGYGSP 943

```

Db 1845 ---VIGRVELPRIML-----PHLLPDTITLCKNEELGFSLCGGHDL 1885
 QY 944 HGDLPYKTVFAKGAASDGRKRGDOIIAVNGSLSEVTHEA-----VAIIKR 994
 Db 1886 Y--QVYISDIPRVAIAEGNLOLDLVHYVNGVSTQGMTELEVRALDMSLPISYLKA 1943
 QY 995 TKGVITLM 1002
 Db 1944 TRNDLPV 1951

RESULT 3 PCT-US94-09943-2

Sequence 2, Application PC/TUS9409943
 GENERAL INFORMATION:
 APPLICANT:
 APPLICANT:
 APPLICANT:
 APPLICANT:
 APPLICANT:
 TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL
 TITLE OF INVENTION: EXPRESSION OF NUCLEOTIDE SEQUENCES FOR NOVEL PROTEIN
 TITLE OF INVENTION: TYROSINE PHOSPHATASES
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 STREET: 600 ATLANTIC AVENUE
 CITY: BOSTON
 STATE: MASSACHUSETTS
 COUNTRY: USA
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/09943
 FILING DATE: 01-SEP-1994
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/115,573
 FILING DATE: 01-SEP-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: TWOMEY, MICHAEL J.
 REGISTRATION NUMBER: P-38,349
 REFERENCE/DOCKET NUMBER: L0461/700000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/720-3500
 TELEFAX: 617/720-2441
 TELEX: 92-1742 EZEKIEL
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2466 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US94-09943-2

Query Match 9.8%; Score 500.5; DB 5; Length 2466;
 Best Local Similarity 22.2%; Pred. No. 6.2e-34;
 Matches 255; Conservative 146; Mismatches 396; Indels 351; Gaps 41;

QY 142 ELQNTAAYSMNPNRYVELMREPSKSLGISTVGRGKMSRLSNG----- 184
 Db 866 QLOMAROSNODADIERASFSRLNLOAESVGFNMGRASITGLASSTLNKLAVPLSV 927
 QY 185 ---EWKGRITIKHY-----LEDSPAGKNGTL-----KPGD--RIEAPSOSESSEPERKPLDS 231
 Db 928 QAEILKRLKSCSELSTYQPLQNSKSKENKRAKSWEEKPREMSKSYHDLQASLYPHRRNVIV 987
 QY 232 VPPPPSAFAEM-GSDHTOSSAS-----KISQDVDKED 263

Db 988 NMEPPQVYAEVLGKPSHQMSRSDAESLAGVTKLNNSKSVASLNSPERKKEHSDSSSIE 1047
 QY 264 EFG---YSKKNIERYGTLTG---ELHMIIELEK-GHSGGLSLAGKNDNR--MSVFIYGI 315
 Db 1048 DGOAYVLDVLRKMSIYSSPEREITLVNLKRDARYGGLGFOITIGEEKMRDLGIFISSV 1107
 QY 316 DNGAAGKGRQIADDELLEINGQILYGRSHONASSIIRKCAKSKYIIFIRKNDVNOA 375
 Db 1108 AFGPRDFHGLKPGDRLIVNSVSLSEGVSHAAIEITLONAPEDVTLVYQREKISKVP 1167
 QY 376 VCPGNAVEPLPNSSENLOKETEPTVTTSDAAVLDLSFEKN-----VOHLLEPK--DOCG 427
 Db 1168 STPVH---LTNEMKNYMKSS---YMGDSALIDSSSKOHMSRGLRIRISNSRGPSCG 1219
 QY 428 L--GIAISEEDTLGVIYIKSLTEHGYAATDGLKYGDDQILAVALDELIVYGPTEKIFSLK 485
 Db 1220 LREGSLSSQDSRTEASLSQSQVNGFFAS---HLGDQW---QESQHSPPSPYIS--- 1269
 QY 486 TAKMTYKLTIIHAENPDSQVPSAAGAAGEKK-----NSSQSLVPOGSPPEPS 535
 Db 1270 --KATEKETFTDSNOSKTKKPGISDVTYSDRGSDMDNATYSSSDHQTPO---ESS 1324
 QY 536 INTSRSTPALFASDPATCPPIPCETIIEISKRTGLSLIVGSDTL--GAFIIE 593
 Db 1325 SVNTSKMKNKTFSSPPK---PGDIFVELAKNDNSIGISVTGQVNTSVRHGIYVKA 1380
 QY 594 YVEEGAACKDGLMAGDQILEVNGIDLKATIDEALINVRQTPORVRLTLVDEAPYKE 653
 Db 1381 VIPOGAASDGRIRHKGRVLAANGVSLGEGATHKAVETLRNGOVVHLLLEGQSTPKE 1440
 QY 654 EV-----C-----DTLIELOKKRKGIGLS--- 674
 Db 1441 HVPVTPQCTLSQONAGQPEKVKTKTQYKDYSPYTEENTFEVKLFKN--SSGLGFSFNR 1499
 QY 675 ---YVGRNDGVFVSDIVYKGIADPGRLLGQDQILLVNGEDVRNAGDAVAALLKCSL 731
 Db 1500 DNILIEQINASIVRYKILPAGOPAAESGKIDVDVILKNGASLKGLSQOEYIALRGTA 1559
 QY 732 GTVTEVGRIRKG-----PPHSF----- 749
 Db 1560 PEVFLLCRPPEVULPEIDTALLTPQSPAQVLPNSSKDSQSPSCVEOSTSDENEMSDK 1619
 QY 750 ---RRPSOTQOVSEGS-----LSSETF-----PL 770
 Db 1620 SKKQCKSPSRBRSYSDSSSGSGEDDLVTAPANISNWSMALLQOTLSNMYSQASHHEARK 1679
 QY 771 SSSST-----SESLESSK----- 784
 Db 1680 SOEDTICTMFFYYPQIPKPEFEDSNPSPLPPDMAFGQSYQPOSESASSSSMDKYHIHI 1739
 QY 785 ---KNLASIIQGLR---TYEMKKGPTISIGISAG---VGSPLGDVP 824
 Db 1740 SEPTROENNTPLKNDLENLEPELEVEELLITLILSEKASISLPTVYKGNORIGCYVHDV- 1798
 QY 825 IFIAMHPTGVAAGTQKLVGDRIYITCGSTEGMTHQAVNLKKNASISIEQVAVAGD 884
 Db 1799 ---IODP---AKSGRRLKPGDRLIKVNDTDTNMHTAVNLDRAASTVNL----- 1844
 QY 885 VSVYTGHOE-PASSLSLTGTLSTISIFQDDLGPPCKSKITLERGPDGLGFSIVGYSGP 943
 Db 1845 ---VIGRVELPRIML-----PHLLPDTITLCKNEELGFSLCGGHDL 1885
 QY 944 HGDLPYKTVFAKGAASDGRKRGDOIIAVNGSLSEVTHEA-----VAIIKR 994
 Db 1886 Y--QVYISDIPRVAIAEGNLOLDLVHYVNGVSTQGMTELEVRALDMSLPISYLKA 1943
 QY 995 TKGVITLM 1002
 Db 1944 TRNDLPV 1951

RESULT 4
 US-08-596-291-3

```

: Sequence 3, Application US/08596291
: Patent No. 5821075
: GENERAL INFORMATION:
: APPLICANT: GONZ, LEONEL JORGE
: APPLICANT: SARAS, JAN
: APPLICANT: CLAESSON-WELSH, LENA
: APPLICANT: HELDIN, CARL-HENRIK
: TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL
: TITLE OF INVENTION: EXPRESSION OF NUCLEOTIDE SEQUENCES FOR NOVEL PROTEIN
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: WOLF, GREENFIELD & SACKS, P. C.
: STREET: 600 ATLANTIC AVENUE
: CITY: BOSTON
: STATE: MASSACHUSETTS
: COUNTRY: USA
: ZIP: 02210
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/596,291
: FILING DATE: 09-AUG-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/115,573
: FILING DATE: 01-SEP-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: GATES, EDWARD R.
: REGISTRATION NUMBER: 31,616
: REFERENCE/DOCKET NUMBER: L0461/7000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617/720-3500
: TELEFAX: 617/720-2441
: TELEX: 92-1742 EZEKTEL
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2465 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-596-291-3

```

Query Match 9.88; Score 498; DB 2; Length 2465;
 Best Local Similarity 21.98; Pred. No. 1e-33;
 Matches 252; Conservative 150; Mismatches 389; Indels 362; Gaps 41;

```

QY 142 ELQMTAVSNMNPQPRVELMREPSKSLGISIVGGRGMSRLSNG-----184
DB 868 QLOQRARNSNDADIDIEASFRSLNLAQESVRGRTMGRAISTGSLASLTAKLAVRLSLV 927
QY 185 --EVMRGIFIRHV-----LEDSPAGKNGTL---KPGD--RIVEAPSOSESEPEKAPLCS 231
DB 928 QAEILKRLSCSELSTLYOPLQNSKSKNDKASKMEKPREMSKSYHDLQASLYPHRKNIY 987
QY 232 VPPPPSAFAM-GSDHQSAS-----KISDDVQKED 263
DB 988 NMEPPQVVALVGRKPSQMSRSDAESLAGVTKLNSKSVASLNSKSPRRKHESDSSIE 1047
QY 264 EFG--YSKNIIRERYGLTG---ELHMIELRK-GHSGIGLS-LAGNKDRSMYSFIYID 316
DB 1048 DPGQAVYVDLHKRMSIYSSPERETTLVNLKKDKAYGIGRPIIIGEKMETDGLGFISSVA 1107
QY 317 PNGAAGKDRLOLADLELEINGQILYGRSHOMASSIICAPSKYKIIIFIRNKDAVNOMAV 376
DB 1108 PGGADEFGCKLPDDRILSVNSVSLGEGVSHAAILIQLNADEVDLTVISQPKERIKSVPS 1167
QY 377 CPGANVRLPSNSENLOKKEPTVYTSDAVDLSSPKN-----VQHLLEPK--DQGL 428

```

```

DB 1168 TPVH-----LTNEMKNYMKSS-----YMODSAIDSSKDHMSRGLRLHISENSFGPSGGL 1219
QY 429 --GIAISEEDLSGVIIKSLLEHGVAANTDGRKXGDDOILANVDDEIVYCIPIEKRIISLKT 486
DB 1220 REGSLISODSRTESASISQSOVNGFFAS---HLGDTW---QSOHGSPPSYIS-----1268
QY 487 AKMTVKLLTHAENPDSQAVPSAAGAAGEKK-----NSSQSLMVPQSGSPERESI 536
DB 1269 -KATEKEFTDSNSKTKKPPISDVTDSGDSDMDENATYSSQDHPOTQ---ESSSS 1324
QY 537 RNTSRSSTPAIFASDPATCPPIIPCEETIETISKRTGLGISIVGSDTL--GAFITHEV 594
DB 1325 VNTSNKNMFKFFSSPPK---PDIFEEVLAKNDSIGISVTGVTMYVHGGIYVKDV 1380
QY 595 YEEGACKDGRMLAGDOILEVNGIDLRKATDEAINVLRORPQVRLTYVDEAPYEE 654
DB 1381 IPQGAESDGRHKRGDRLVANGVSLGEGATHKQAVETLRNTGQVYVHLLLEKQSPSTKEH 1440
QY 655 V-----C-----DITLIELOKRPKGLS-----674
DB 1441 VPVTPQCTLSQNAQGGQPEKVKTKTQKDYSPYTEENTFEVKLFKN--SSGLGFSFRED 1499
QY 675 --YGRKRDVGFVSDIYKGIADPDGRILGDDOILLVNGDVYRNAGQEAVALKCSLG 732
DB 1500 NLPEQINASIVRYKKLFGQPAESGKIDVDVILKVNGLSLGISOEVISALRGTA 1559
QY 733 TVTLEVGRIKAG-----PFHSE-----749
DB 1560 EVFLLICRPPGVLEIDTALTPLQSPAQVLRPNSSKDSQPCVEOSTSDENEMSDKS 1619
QY 750 ---RRPSQTSQVSEGS-----LSFTF-----PLS 771
DB 1620 KKQCKSPSRSDSYSDSGEDDLVTAPANISNSTWSALHOTLSNNVSOAQSHHEAPKS 1679
QY 772 GSST-----SELSSSK-----784
DB 1680 QEDTICTMFTYPOKIPNKPEBEDSNPSPLPDPMAPOGSIPOSSASSSMDKHHIHS 1739
QY 785 -----KNALASEIOGLR-----TYEMKGPYDLSISYAG--VSPPLGDPVI 825
DB 1740 EPTROENMTPLKNDLENLEDFELEVELLITLISEKASLGFTYTKGNRGCVVHV--1797
QY 826 FIAMHPTGVAQTOQLRVDGRIVTICGSTEGTTHPOAVNLKNASGSIEM-----QV 879
DB 1798 ---IODP---AKSDGRILKPGRLIKVNDTQVNTHTDAVNLRAASKTVRLVYGRPRI 1851
QY 880 VAGDGV--SVYTGHHQEPASSLSFTGLTSTISFPQDGLGPPQCKSITLERGPDGFSYNG 938
DB 1852 TQNTNVASFATGH-----KLTCKNEELGFSLCG 1879
QY 939 GYGSPPHGLPIYVTVFAKGAASEDRGLRKQDQIIAVNGSLGCVTHDEA-----V 989
DB 1880 GHSDLV--QVYISDINPRVAALIEGNQLDDVHYVNGVSTQGMTELEVNRALDMSLPS 1937
QY 990 AILKRTKGTVM 1002
DB 1938 LVLKATRDLPVV 1950

```

RESULT 5
 US-09-100-804-3
 Sequence 3, Application US/09100804
 Patent No. 6066472
 GENERAL INFORMATION:
 APPLICANT: GONZ, LEONEL JORGE
 APPLICANT: SARAS, JAN
 APPLICANT: CLAESSON-WELSH, LENA
 APPLICANT: HELDIN, CARL-HENRIK
 TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL
 TITLE OF INVENTION: EXPRESSION OF NUCLEOTIDE SEQUENCES FOR NOVEL PROTEIN
 NUMBER OF SEQUENCES: 34
 CORRESPONDENCE ADDRESS:

ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
 STREET: 600 ATLANTIC AVENUE
 CITY: BOSTON
 STATE: MASSACHUSETTS
 COUNTRY: USA
 ZIP: 02210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/100,804
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/596,291
 FILING DATE: 09-AUG-1996
 APPLICATION NUMBER: US 08/115,573
 FILING DATE: 01-SEP-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/09943
 FILING DATE: 01-SEP-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: GATES, EDWARD R.
 REGISTRATION NUMBER: 31,616
 REFERENCE/DOCKET NUMBER: L0461/7003
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441
 TELEX:
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS: 3
 LENGTH: 2465 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-100-804-3

Query Match 9.8%; Score 498; DB 3; Length 2465;
 Best Local Similarity 21.9%; Pred. No. 1e-33;
 Matches 252; Conservative 150; Mismatches 389; Indels 362; Gaps 41;

DB 142 ELQNTAYSNWNOPRVLELREPSKSLGISTVGRGMSRLSNG----- 184
 DB 866 QLOMRAROSNODADIERASFRSLQAESVGFNMGRATISTGSLASTYLINKLAVAPLSV 927
 QY 185 --EVMRGITIKHY-----LEDSPPAGKNGTL---KPGD--RIVEAPSOSESSEPEKAPLCS 231
 DB 928 QAEILRLKRLSCSELSTYQPLQNSKSKENKAKSWEEKPREMSKSYHDLQASLYPHRKNVIV 987
 QY 232 VPPPPSAFAEM-GSDHTOSSAS-----KISODVDEKED 263
 DB 988 NNEPPQYVAELVGKPSHQMSRSDAESLAGVTKLNKSKSVASLNRPERKRIHESDSSSTE 1047
 QY 264 EFG--YSKNIRERYGTLTG---ELHMITLEK-GHSGLGLS--LAGNKRDSRMSVFTVGID 316
 DB 1048 DGGQAVYLDVLVLRKMSIVSPEREITLVNKKDKAKYGLGFOIIGGKMETDGLGIFISVA 1107
 QY 317 PNGAAGKQGRLOADELLINGOILYGRSHONASSIICAPSKVYKLIIFIRNDVAVNOVAV 376
 DB 1108 PGGPADFHCCLPGRDLISVNSVSLGVSHPAAIETLQNAPEDEVTLVISOPEKISKVPS 1167
 QY 377 CGNNAVEPLPSNSENLOKTEPEVTYVSDAAVDSLFFKN-----VQHLELPK--DQGL 428
 DB 1168 TPVH---LTNEMKNVMKKS---YMODSALDSSSKDHMSNGTLRHSNSNFGSPSGGL 1219
 QY 429 --GIAISEEDTISGVYIKLTERGVAATDGRKLVGQDIIAIVGVPITKEFISLKT 486
 DB 1220 RGSLSSODSRTESASLSOSQVNGFFAS---HLGDQTM--QESQHGSPSPSVIS--- 1268
 QY 487 AKMYVLTITHAENPDSQAVPSAAGAAGEKK-----NSSQSLMVPQSGSPPESTI 536

DB 1269 -KATEKETDSDNOSKTKKPRGSDVTDYSDRSDSDMEATYSSSDHQHPKQ---ESSSS 1324
 QY 537 RNTSSSTPAIRASDPANCPIIPGCEETIEISKRTGLGLSTVGSVDLL--GAPLIHEV 594
 DB 1325 VNTSNKMNFKETSSSPK---PGDIFEEVLAKNDNSLIGSTVGCVNTSVRRGIGYVKKV 1380
 QY 595 YEEGAACKQGRIMADQILEVNGIDLRKATHEAIVNLRQTPORRLTYRDEAPYKEBE 654
 DB 1381 IPQGAESDGRILHKDGRVLAVNGVSLGATHKQAVETLNTQOVVHLLLEKQSPYSKEH 1440
 QY 655 V-----C-----DTLTLEQKPKGGLS----- 674
 DB 1441 VPTPQCCLSDQNAQGGQGEKVKTTQVKDYSFYVEENFEVKLFKN--SSGLGFSRSD 1499
 QY 675 --TVGRNDTVGVSDIYKGIADPDGRLIGDQIILVNGEDVRNASEVAVALLKCSLG 732
 DB 1500 NLIPQINASIYRVKLFAGQPAESGKIDVGVILKVGASLKGISQOEIVISALRGTA 1559
 QY 733 TVTLEVGRIKAG-----PFHSE----- 749
 DB 1560 EVFLLCRPPGVLPETIDTALLTPQSPAQVLPNSSKSSQSPCYVEOSTSSDBENKSDKS 1619
 QY 750 ---RRPQTSQVSEGS-----LSSTF-----PLS 771
 DB 1620 KKQCKSPSRSDYSOSSSGGEDDLYTAPANISNSTWSALHOTLSNMVSOAQSHHAPKS 1679
 QY 772 GSST-----SESLSSSK----- 784
 DB 1680 QEDTICTMFEYPOKIPNKEFEDSNPSPLPDPMAQSOYOPSESASSMKYHIHHS 1739
 QY 785 -----KNALASEIQLR---TVEMKKGTDSLGISING---VSPGLDVP 825
 DB 1740 EPTROENMTPLNNDENHLEDELEVELLITLIKSEKASLGFTYKGNORICYVADV-- 1797
 QY 826 FLAMHPTGVAQOTOKLRYGDRIVITIGTSTEGMTHTQAVNLKNAQSIEM-----QV 879
 DB 1798 ---IQDP---AKSDGLKAGDRILKYNPDVYVNMHTDANVLLRAASKYRVLVIGVPR 1851
 QY 880 VAGGV-SVVTGHOBEPASSLSFTGLSTSLFPDDLGPPCKSTLERGPDGLGFSYV 938
 DB 1852 TONTNVAEPATCH-----KLTCKNKEBLGFSLCG 1879
 QY 939 GYGSPHGDLPYKTYFAAGASSEDGRKRGDIIAVNGOSLEGYTHEA-----V 989
 DB 1880 GHDSTL--QVYVTSIDNIPRSVAITEGNQLDVLVHYVNGVSTQGMTELEVNRALDMSLPS 1937
 QY 990 AILKRTKGTVTLM 1002
 DB 1938 LVLKATRNDLPVY 1950

RESULT 6
 US-09-045-632-2
 Sequence 2, Application us/09045632
 Patent No. 6001575
 GENERAL INFORMATION:
 APPLICANT: Huganir, Richard L.
 APPLICANT: Dong, Hualing
 TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
 TITLE OF INVENTION: GRIP-RELATED MOLECULES
 NUMBER OF SEQUENCES: 105
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
 STREET: 130 Water Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,632
FILING DATE: 19-MAR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,016
FILING DATE: 19-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Corless, Peter F.
REGISTRATION NUMBER: 33,860
REFERENCE/DOCKET NUMBER: 48147/1699-CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1112 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-045-632-2

```

```

Query Match      6.7%; Score 340; DB 3; Length 1112;
Best Local Similarity 22.6%; Pred. No. 1.5e-20;
Matches 173; Conservative 90; Mismatches 265; Indels 238; Gaps 22;

```

```

QY 413 EKNVHLEPRDQ-GIGIAISEDTLSGV-----IKSLTEHVAATGRKLVQDI 464
DB 48 FKSGTVELMKREGTTLC-----TVSGIDKDKPRVSNLRGGIARSDQLDVDYI 101
QY 465 LAVDEIVGVPIKFKISLTAKMTVKLTTHAENPDQAVPSAAGASGKKNSSQSLM 524
DB 102 KAVGINLAKRHEIISLTAKNGERVVLEVEYELP----- 137
QY 525 VPQSGSPESIRNTSSRSTPAIRASDPATCPIIPGCTTIEISKGRG--LGSIYGS 582
DB 138 -----PVSIOGSS-----VME-----RTVEVTLHKRGNTFGFVIRGA 170
QY 583 D---TLGAFIHEVYEGACAKDGRMLAGDQILEVNGIDLRKATHDAINVLRQTPQV 639
DB 171 HDNRKSPVITCYRPGCPDRRECTIKPGRILSVDRILGLTHAEMSLIKQCGDEA 230
QY 640 RLTL-YRDEAPYKEEVCDTLTIELQKPKGKGLSIYVK--RNDTGVFVSDIYKGIAD 696
DB 231 TLLEIVDSANDSVATAGPLVEAKTPGASLSGLVALLTSCVKQVIVIDKISASIA 290
QY 697 PDGRLIQGDQILLVNGEDVRNASEAVALLKCSLGTVTLV-----GRK-AGPFH-- 747
DB 291 RCGALHVGDIHLSIDGSMECTLAETQFLGNTTDQVKEILPHHQTRALKGPDHVKI 350
QY 748 -----SERPSQ 754
DB 351 QRSRQLEWDPWASSQCVHTNNHHNPHHPDCHVPALGFPKALLPNSPPMAVSSSS 410
QY 755 TSOVSEGLSSTFPLSGSSTS-----ESLESSKKNALASEIQLGLTV 798
DB 411 MSAYSLSLNNGTLPRLSYSTSPRGTMWRRLKKKDFKSSLSLASYGLAGQVHFFET 470
QY 799 E--KKGPTDLSGISIAGV--GSPLGDPVPIITAMHHPGVAQAQOKLAVGDRIVTIG 855
DB 471 EVVLADPVYTFGQILOGSVFATELTSSPPLISYIEADSPERCGVLOQIDRVMAIN 530
QY 856 TEGTHQAVNLKKNASG-----IEMQVYAGDVSVYVGHGHEPASSSLSFGLNLS 909
DB 531 TSDTFEBANOLDKSDSTTSKVTLEIEFDVAESVIPSSTGFHVLPKKHVEL-GITIS 589
QY 910 IFQDGLGP-----GCK--- 921
DB 590 PSSRKPDPPLVSIIDIKKGSVAHRTGTLELGDKLLAIDNIRLSDCSMEDAVQILOG 649

```

```

QY 922 -----SITLER--GPDGLGSIYGVGSPHDDLPYVKTFAK 957
DB 650 KLIKRDENSDDEQSSGAILYVELKRYGGPLGITIS-----GTEPFDPIIISLTKG 704
QY 958 GAASEDKLRKGDQIIANVGSLGCVTHEEVAVALIKRKGVTVMV 1003
DB 705 GLAERTGAIHIGDRILAINSSSLKGRPLSEDIHLLQMAGETVTYTKI 750

```

RESULT 7

```

US-09-045-632-3
Sequence 3, Application US/09045632
Patent No. 6001575
GENERAL INFORMATION:
APPLICANT: Huganir, Richard L.
APPLICANT: Dong, Hualing
TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
TITLE OF INVENTION: GRIP-RELATED MOLECULES
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,632
FILING DATE: 19-MAR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,016
FILING DATE: 19-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Corless, Peter F.
REGISTRATION NUMBER: 33,860
REFERENCE/DOCKET NUMBER: 48147/1699-CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1112 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-045-632-3

```

```

Query Match      6.7%; Score 340; DB 3; Length 1112;
Best Local Similarity 22.6%; Pred. No. 1.5e-20;
Matches 173; Conservative 90; Mismatches 265; Indels 238; Gaps 22;

```

```

QY 413 EKNVHLEPRDQ-GIGIAISEDTLSGV-----IKSLTEHVAATGRKLVQDI 464
DB 48 FKSGTVELMKREGTTLC-----TVSGIDKDKPRVSNLRGGIARSDQLDVDYI 101
QY 465 LAVDEIVGVPIKFKISLTAKMTVKLTTHAENPDQAVPSAAGASGKKNSSQSLM 524
DB 102 KAVGINLAKRHEIISLTAKNGERVVLEVEYELP----- 137
QY 525 VPQSGSPESIRNTSSRSTPAIRASDPATCPIIPGCTTIEISKGRG--LGSIYGS 582
DB 138 -----PVSIOGSS-----VME-----RTVEVTLHKRGNTFGFVIRGA 170
QY 583 D---TLGAFIHEVYEGACAKDGRMLAGDQILEVNGIDLRKATHDAINVLRQTPQV 639

```

```

Db 171 HDDRNRKSRPVITCVPRGCGDDREGTIKRGDRLLSVDIGRLTGTHAEMSLKCGQGA 230
Qy 640 RLL-YRDEAPRKEEVEDTLTIELQKRGKGLSTYVK--RNDTGVVSDIYVGGLAD 696
Db 221 TLLIYDVASMDSVATASGPLVEAKTPGASLVALTTSVCCNKQVYIDIKKSASLAD 290
Qy 697 PDGRLIOGQDILLVNGEDVRNASQEAVALLKCSLGTVLEY-----GRK-AGPFH--- 747
Db 291 RGAALHVGDHILSIDGTSMETCTLAETQFLGNTDQVKELELPHQRLALKGDPHVI 350
Qy 748 -----SERPSQ 754
Db 351 QRSDROLPMWPMWASSQSVHTNNHHNPHDHCRVPLGFPKALTPNSPRAWSSSPPTS 410
Qy 755 TSQVEGSLSTFPPLSGSSTS-----ESLESSSKKNAASEIOLGRLV 798
Db 411 MSAYSLSLNMGTLPRLSTSPRGTMMRRRLKKDFKSSLSLASTVGLAQVHTET 470
Qy 799 E--MKKPTDSIGISAGV--GSPLDGVPIFTAMMPTGVAOQOKLVRGDRIVTICGS 855
Db 471 EYVLADPVTGCGIOLQSGVFATETLSPPLISTYIADSPARCGLDIGNVMAINGIP 530
Qy 856 TEGMTHTQAVNLKNASGS-----TEMQVYAGDVSVVTHHOBPASSLSFTGLTSTS 909
Db 531 TFDSTFEENQOLLRDSSTISKVTLIEFDVAESVIPSSGTFHVKLPRKHSVEL-GITISS 589
Qy 910 IFODLGRP-----OCK--- 921
Db 590 PSSRRKGPVYISDIKGSVAHRTGLEGLKLAIDNIRLDCSMEDAVOILQOCEDLV 649
Qy 922 -----SITLER--GPDGLGFSTVGGSGPHDLPYVTVFAK 957
Db 650 KIKIKDENSDEQSSCAIITYVELKRGGLGITIS-----GTEEPDPIIISLTG 704
Qy 958 GAASEDGLRKGDQIIAVNGQSEGVTHEEVAVALIKRTKGVTLV 1003
Db 705 GAERTGAIHIGDRILAINSSSLKGPISLSEDIHLLQMAGETVTLKI 750

```

RESULT 8

```

; US-09-045-632-15
; Sequence 15, Application us/09045632
; Patent No. 6001575
; GENERAL INFORMATION:
; APPLICANT: Hugenit, Richard L.
; APPLICANT: Dong, Hualing
; TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
; NUMBER OF SEQUENCES: 105
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/045.632
; FILING DATE: 19-MAR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/041,016
; FILING DATE: 19-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Corleiss, Peter F.
; REGISTRATION NUMBER: 33,860
; REFERENCE/DOCKET NUMBER: 48147/1699-CIP
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 702 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-045-632-15

```

```

Query Match 6.6%; Score 335; DB 3; Length 702;
Best Local Similarity 22.5%; Pred No. 1,9e-20;
Matches 171; Conservative 90; Mismatches 261; Indels 238; Gaps 22;

```

```

Qy 419 LELPRDQ-GIGIAISEEDTLGVI-----IKSLTEHVATDGRKLVGQDIIAVNDE 470
Db 3 VELMKKEGTLLGC-----TVSGGIDKQKPRVSNLRGGIARSDQDLVDGYIKAVNGI 56
Qy 471 IYVGPRIEKFISLLTAKKTVKLTTHAENPDQAVPSAAGAAGEKKNSQSLAMPQSGS 530
Db 57 NLAKEFRHEIISLLKNVGERVYLEVEYELP----- 86
Qy 531 PEPEIRNRSRSTPAIFASDPATCPRIIPGCTETIISKGRG--LGLSTVGGSD--TL 585
Db 87 --PVSTQSS-----VME-----RTVEVTLHKGMTGFEYIRGGAHDRNK 125
Qy 586 LGAFITIEYEEGCAKQDRLWAGDOILEVNGIDLAKATHDEAIVNLTPOVRVRLT-Y 644
Db 126 SRPVYITCRPGPDRETIKRGDRLLSVDIGRLTGTHAEMSLKCGGAEATLILEY 185
Qy 645 RDEAPRKEEVEDTLTIELQKRGKGLSTYVK--RNDTGVVSDIYVGGLADPDGRIL 702
Db 186 DVSANDSVATASGPLVEAKTPGASLVALTTSVCCNKQVYIDIKKSASLADRGALH 245
Qy 703 QGDQIILLVNGEDVRNASQEAVALLKCSLGTVLEY-----GRK-AGPFH--- 747
Db 246 VGDHILSIDGTSMETCTLAETQFLGNTDQVKELELPHQRLALKGDPHVIKIQRSBQ 305
Qy 748 -----SERPSQTSQVSE 760
Db 306 LPMWPMWASSQSVHTNNHHNPHDHCRVPLGFPKALTPNSPRAWSSSPTSMSAYSL 365
Qy 761 GSLSTFPPLSGSSTS-----ESLESSSKKNAASEIOLGRLTVE--MKK 802
Db 366 SSLNMGTLPRSLSTSPRGTMMRRRLKKDFKSSLSLASTVGLAQVHTETTEVTLTA 425
Qy 803 GPTDSIGISAGV--GSPLDGVPIFTAMMPTGVAOQOKLVRGDRIVTICSTEGMTH 861
Db 426 DVTGFGIOLQSGVFATETLSPPLISTYIADSPARCGLDIGNVMAINGIPEDSTF 485
Qy 862 TQAVNLKNASGS-----TEMQVYAGDVSVVTHHOBPASSLSFTGLTSTSIFODL 915
Db 486 EKANOLLRDSSTISKVTLIEFDVAESVIPSSGTFHVKLPRKHSVEL-GITISSPSRKP 544
Qy 916 GPP-----OCK--- 921
Db 545 GQPLVYISDIKGSVAHRTGLEGLKLAIDNIRLDCSMEDAVOILQOCEDLVKIKRK 604
Qy 922 -----SITLER--GPDGLGFSTVGGSGPHDLPYVTVFAKGAASED 963
Db 605 DEDNSDEQSSCAIITYVELKRGGLGITIS-----GTEEPDPIIISLTGKGLAENT 659
Qy 964 GRLKRGDQIIAVNGQSEGVTHEEVAVALIKRTKGVTLV 1003
Db 660 GAIHIGDRILAINSSSLKGPISLSEDIHLLQMAGETVTLKI 699

```

RESULT 9

```

; US-09-045-632-16
; Sequence 16, Application us/09045632
; Patent No. 6001575

```

GENERAL INFORMATION:
APPLICANT: Huganir, Richard L.
TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,632
FILING DATE: 19-MAR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,016
FILING DATE: 19-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Corless, Peter F.
REGISTRATION NUMBER: 33,860
REFERENCE/DOCKET NUMBER: 48147/1699-CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1018 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-045-632-16

Query Match 6.6%; Score 335; DB 3; Length 1018;
Best Local Similarity 22.5%; Pred. No. 3.6e-20;
Matches 171; Conservative 90; Mismatches 261; Indels 238; Gaps 22;
QY 419 LELPKDGG-GIGIAIASEDDLSGIY-----IKSLTEHGVAATDGRKLVQDOLAVDE 470
DB 3 VELKKKSTLTGCG-----TVSSGIDKDGKPRVSNLRQGGIAASDDLDVGDYIKAVNGI 56
QY 471 IVVGYPLEKFTSLTKAKMTVKLTIIHAENPDQAVPSNAGAAGEKKNSQSLWVPOGS 530
DB 57 NLAQFRHDEILSLKKNYGERVYLEVEYELP----- 86
QY 531 PEPISIRNTSSSSPAIFASDPATCPIIPGCEETIETISKGTG--LGISIVGSD--TL 585
DB 87 --PVSIGSS-----VME-----RIVEVTLHKEGWTGFEVIGAGHDDRNK 125
QY 586 LGATIIHEVVEGSAACGRIMAGDOLLEVNGIDLKRAIHEANLVNROTPQRVRLTL-Y 644
DB 126 SRPVVITCVPRGGPDDEGTIKPGDRILSVGIRLLGTTIHAANSILKQCCQEAFTLLEY 185
QY 645 RDEAPYKEEVECDLTLELQKPKGKGLSLVGR--RNDTGVFSDIYKGIADPDGRILI 702
DB 186 DVSAMDVAATASGPLVLEVAKTPGASIGVATTSVCCNKQVIYDKIKSASIAIRCGLH 245
QY 703 QGDOILLVNGEDVARNASQEAVALLKCSLGTVTLEV-----GRIK-AGPEH----- 747
DB 246 VGDHILSIDGSMETCTLAETQFLGNTTPOVKLEILPHHQTRIALKGPDRHVKIORSDRQ 305
QY 748 -----SERPSTSOVSE 760
DB 306 LPMDFMASSQCSVHTNHNHNPDRHCHRVPALEGPKALTPNSPAMVSSSPSTMSAYSL 365

QY 761 GLSSFTPELSSSSNS-----ESLESSKKNALASEIQLRTVE--MKK 802
DB 366 SLMNGKITPRSLYSTSPGCTMMRRRLKKDKRKSLSLASSIVGLAGOVYHETTEVYVLA 425
QY 803 GPTDSLGISIAGV-GSPPLGDPVPIFIAMHPTGVAQAOTKLRYDRIYITIGSTEGNTH 861
DB 426 DPVGVFGIQLGVSFATETLSSPPLISYIEADSPAEKRGVLIQIDRVVAINGIFTEDSTF 485
QY 862 TQAVNLKNAAGS-----IMQYVAGDVSVYGHQEPASSLSLFTSTIFPDDL 915
DB 486 EEAQQLNDSSITSKVTLEIFEDVAESVIPSSGTFHVKLPKRHVEL-GITISPPSSRRP 544
QY 916 GPP-----QCK----- 921
DB 545 GDPPLVSDIKKGSVAHRTGTLELDKLLAINDIRLDSQSMEDAVOILQCCEDYKLRK 604
QY 922 -----SITLER--GPDGLGFSIVGSGSPHGDLPYVKTVFAKGAASED 963
DB 605 DEKNSDEQESSGAILIYVELKRYGPGILITIS-----GTEEPFPIIISLTKGGLAERT 659
QY 964 GLKRGDQIIAVNGQSLGVTHEEVAVALKRTKGTVTLMV 1003
DB 660 GAHIGDRILAINSSSLKGPPLSEDIHLLQAGETVYTKI 699
RESULT 10
US-09-045-632-32
Sequence 32, Application US/09045632
Patent No. 6001575
GENERAL INFORMATION:
APPLICANT: Huganir, Richard L.
TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,632
FILING DATE: 19-MAR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,016
FILING DATE: 19-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Corless, Peter F.
REGISTRATION NUMBER: 33,860
REFERENCE/DOCKET NUMBER: 48147/1699-CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 1061 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-045-632-32
Query Match 6.6%; Score 335; DB 3; Length 1061;

Best Local Similarity 22.5%, Pred. No. 3.9e-20;
Matches 171; Conservative 90; Mismatches 261; Indels 238; Gaps 22;

```

OY 419 LELPKDQG-GIGIAISEEDTLGVT-----IKSLTEHGVATDGRLLKVGDOIILAVDDE 470
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3 VELMKKEGTTGCC-----TVSGIDKDKGKPRVSNLROGDIARSDQDLDVGDYIKAVNGI 56
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 471 IYVGPRIEKFISLLTAKTKVTLTHAENPDSQAVPSAAGASGEEKKSSQSLMPQSS 530
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 57 NLAKEFRHEIISLLKNVGERVLEVELELP----- 86
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 531 PEPEIRMTSRSSSTPAIFASDPATCPRIIPGCTETIEISKRTG--LGLSTVGSD---TL 585
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 87 --PVSIOGSS-----VMF-----RIYEVTLAKKGNFEGFVIRGADHDRK 125
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 586 LGAFITHEYEYEGCAKDGRLMAGDOIILEVNGIDLRKATHDEAIVLROTPQVRVLT-LY 644
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 126 SRPVYTCVRPGPDREGTIRKDRLLSVDDGIRLLGTHAEMSIKQCGGEATLLLEY 185
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 645 RDEATYKKEEVCDLTLELOKKRPGKGLSTYVK--RNDTGVFVSDIYKGIADPDGRLI 702
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 186 DVSAMDVAATASGPLIVEAKTPGASLVALTTSVCCNKQVIVIDIKISASADIACGALH 245
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 703 OGDOIILVNGEDVNRASQEAVALKCSLGVTLVY----GRIK-AGPFH----- 747
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 246 VGDHLISIDGTMEYCTLAETQFLGNTTDQYKLELPHQTRALKGPDHAKIGRSDQ 305
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 748 -----SERRPSQTSQVSE 760
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 306 LPMDWASSQCSVHTNNHHNPHRPHCRVPLGFPKALTPNSPRAWSSSSPTMSAYSL 365
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 761 GLSLFTPLRSGSSTS-----ESLESSKKNALASEIOGLTVE--MKK 802
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 366 SSLNMGCTLPRLSYSTSPCTMMRRKKKDFKSSLSLASTVGLAGQVHTTETTEVLTA 425
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 803 GPTDLGISIAGV--GSPGLDVPIFIAMNHPTGVAQTKLRVGRIVITICSTEGMTH 861
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 426 DVTGFEGLDQSVATEHTLSSPRLISYLEADSPAREGCVLIDGRVMAINIPREDSTF 485
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 862 TOAVNLLKNAGS-----TEMQVYAGGVSVYTGHHQEPASSLSFTGLSTSTFODLL 915
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 486 EKANOLLRDSSITSKYVLEIEEDVAESVTPSSGTFPHVKLPKHSVEL-GITISPSRSRP 544
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 916 GPP-----OCK----- 921
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 545 GCPPLVITDKKKSVAHRTGTLELGDKLAIIDNIRLDSCEMEDAVOILQCCEDLVKLIKTK 604
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 922 -----STILER--GPDGLGFSIVGGYSGSPHGDLPYVKTVEPAKGAASED 963
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 605 DEDNDEQESSGAIITVETLKRKGGLGITIS-----GTEEPFDPIIISLTKGGLAERT 659
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 964 GRLKRGDOIIVANGOSLSEGVTHEEVAAILKTKKGVITLMV 1003
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 660 GAHIGDRILAINSSSLKKGKPLSEDIHLLQMAGETVTLTKI 699
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 11
US-08-410-804-1

; Sequence 1, Application US/08410804
; Patent No. 5632994
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Sato, Takaki
; TITLE OF INVENTION: FAS ASSOCIATED PROTEINS
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cathryn Campbell
; STREET: 4370 La Jolla Village Drive, Ste 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/410.804
FILING DATE: 27-MAR-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/259,514
FILING DATE: 14-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1389
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-410-804-1

Query Match 6.6%, Score 334.5; DB 1; Length 610;
Best Local Similarity 23.0%, Pred. No. 1.7e-20;
Matches 155; Conservative 72; Mismatches 221; Indels 225; Gaps 21;

```

OY 474 GYPIKFKISLLTAKTKVTLTHAENPDSQAVPSAAGASGKK-----NSSQSL 523
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2 GSPSPSVIS-----RATEKETFTDSNQSKTKRPGISDVTDYSDRQSDMDERTYSSSDH 56
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 524 WPOSQSPESIRMTSRSSSTPAIFASDPATCPRIIPGCTETIEISKRTGGLSTY---- 579
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 57 QTPKQ---ESSSVTSMNMNMNFTSSSPK---PGDIEVELAKNDNSLGISTVLEFD 109
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 580 -GGSOTLL--GAFITHEYEYEGCAKDGRLMAGDOIILEVNGIDLRKATHDEAIVLROTP 636
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 110 KGVMTSVRHGGIYVKAIVPQGAASDGRIRHKGDRVLAVNGVSLGATGHQAVETLRMTG 169
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 637 QWRVLTLYDEAPYKKEEV---C-----DTLTI 661
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 170 QVHLLLEKQSPTSKEHVPVTPQCTLSDQNAOGGPEKVKKTTQVKDYSFVTEETFEV 229
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 662 ELQKRPKGIGLS-----IVKRNDTGVFVSDIYKGIADPDGRLIOGDOIILVNGED 714
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 230 KLFKN--SSGLGFSFRREDNLPEQINASTIVRYKKLFPGQPALESKIDVGLKVNAGS 288
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 715 VNASQEAVALKCSLGLTVLEVGRIKAGP-----HSERRPSQT 755
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 289 LKGLSQOEIVSALRGTAPEVFLILCRPPGVLPETIDTALLTPQSPAVLPNSSKDSQP 348
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 756 SQVSEGSLSFTPLSGSTSESSSKKNALASEIOGLRVEMKKGTDSLGI SIAGG 815
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 349 SCVEQ-----STSSDENEMDSKKQC-----KSSRDSYSDSG 384
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 816 VSPPLGDVPIFIAMNHPTGVAQTKLRVGRIVITICSTEGMTHTOAVNLLKNAGSI 875
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 385 SGE-----DDLVI-----APAINISSTWS 404
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 876 EMQVYAGGVSVYTGHHQEPASSLSFTGL-----TSYSIFODDLP--- 917
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 405 ALHQTLNNVSOAGSHHEAPKQOEDTICTMFEYPOKIPKREFEDSNPSPLPPDAFGS 464
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 918 --POCKS-----TLERGPDG 931
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 465 YOPQSESASSMDXYHHHISEPPEOENWPLKNDLENHLEDFELEVLLTTLIKSEK 524
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 932 -LGFSIVGGYSGSPHGDLPYVKTVEPAKGAASEDGRLLKRGDOIIVANGOSLSEGVTHEEVA 990
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 535 SLGFTVTKG---NORIGCYHVDI-QDPAKSDGRLKRGDLIKVNDVTNMTHTDAVN 579

QY 991 ILKRTKGTVTLAV 1003

Db 580 LLRAASKTVRLVI 592

RESULT 12

US-08-259-514-1

Sequence 1, Application US/08259514

Patent No. 5747245

GENERAL INFORMATION:

APPLICANT: Reed, John C.

APPLICANT: Sato, Takaaki

TITLE OF INVENTION: FAS ASSOCIATED PROTEINS

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cathryn Campbell

STREET: 4370 La Jolla Village Drive, Ste 700

CITY: San Diego

STATE: California

COUNTRY: United States

ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/259,514

FILING DATE: 14-JUN-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn

REGISTRATION NUMBER: 31,815

REFERENCE/DOCKET NUMBER: P-LJ 9954

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 535-9001

TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 610 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-259-514-1

Query Match 6.68; Score 334.5; DB 1; Length 610;

Best Local Similarity 23.08; Pred. No. 1.7e-20;

Matches 155; Conservative 72; Mismatches 221; Indels 225; Gaps 21;

QY 474 GYPIEKFTSLTKAKMTVKLTIIHAENPDQAVPAAGAAGEK-----NSSQSL 523

Db 2 GSPSPVIS-----KATEKEFTDSNQSRTKPKPISDVTDSGDSDMDATYSSSDH 56

QY 524 MVPOSGSPSPSINTSRSSPAIFASDPATCPIIPGCEETIETSKGTGLGISIV----- 579

Db 57 QTPKQ---ESSSVNTSNKNMFKTFSSPPK---PGDIFEVELAKNDNSLGISVTVLFD 109

QY 580 -GGSDTL--GAFIIHEYEGACKDGRLMAGDQILLEVNGIDLRKATHDAIVLQTP 636

Db 110 KGVNTSVRRHGIVYKAVIPQGAESDGRJHKGRVLAVNSVLEGATHKQAVETLRNTG 169

QY 637 QRVALTLYRDAPYKEEV-----C-----DTLTI 661

Db 170 QVAVLLLEKGGSPSKSHVPTPCQTLSDQNAOGSPKVKTKTQVNDYSVTENTFEV 229

QY 662 ELQKKPKGLGLS-----IVGRNDTGVFVSDIVKGLADPDGRLIQDQILLVNGED 714

Db 230 KLFKN-SSGIGFSSREDNLPQIINASIYVKKLPFGQPAESGKIDVGVILKVGAS 288

QY 715 VRNASQAVALLKCKSLGTVLEVGRIKAGF-----HSRRRSQT 755

Db 289 LKGLSQDEVISALROTAPEVLLLCRPPGVLPEDITALLTPLOSPACVLPNSSKSSQP 348

QY 756 SOVSEGLSFTPLSGSSSTSESSSKKNALNSIQLGTLVEKKKQPTDSLGISING 815

Db 349 SCVED-----STSDENEMSKSKOC-----KSPSRDSTSDSG 384

QY 816 VSPGLDVPFIAMHPTGVAQOTOKLRGDRIVTICSTEGMTHQAVNLKNASGSI 875

Db 385 SGE-----DDIVT-----APANISSTWSS 404

QY 876 EMQVAGDVSVTGHQEPASSLSFTGL-----TSTSIQDGLP--- 917

Db 405 ALHQTLSNMVSOAOSHREAPKSOEDTICTMYYPQKIPNKPFEFEDSNPPLPDPAPQS 464

QY 918 --POCKS-----ITLERGPDG 931

Db 465 YQPSSESASSMDKYHIHISEPTROENWTPLNKDLENHLEDEVELLITLIKSEK 524

QY 932 -LGFSTVGYGSPHDDLPIYKTVFAKGAASEDGRKRGDQIIAVNGSLGVTHEEAVA 990

Db 525 SLGFTVTKG---NORIGCYHVDI-QDPAKSDGRLKRGDLIKVNDVTNMTHTDAVN 579

QY 991 ILKRTKGTVTLAV 1003

Db 580 LLRAASKTVRLVI 592

RESULT 13

US-08-858-311-1

Sequence 1, Application US/08858311

Patent No. 5876939

GENERAL INFORMATION:

APPLICANT: Reed, John C.

APPLICANT: Sato, Takaaki

TITLE OF INVENTION: FAS ASSOCIATED PROTEINS

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cathryn Campbell

STREET: 4370 La Jolla Village Drive, Ste 700

CITY: San Diego

STATE: California

COUNTRY: United States

ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/858,311

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/410,804

FILING DATE: 27-MAR-1995

APPLICATION NUMBER: US 08/259,514

FILING DATE: 14-JUN-1994

ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn

REGISTRATION NUMBER: 31,815

REFERENCE/DOCKET NUMBER: P-LJ 1389

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 535-9001

TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 610 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-858-311-1

Query Match 6.6%; Score 334.5; DB 2; Length 610;
 Best Local Similarity 23.0%; Pred. No. 1,7e-20;
 Matches 155; Conservative 72; Mismatches 221; Indels 225; Gaps 21;

QY 474 GYPIKFTSLTAKMTVLTTHAENPDSQAAPSAAGSGERK-----NSSQSL 523
 DB 2 GSPSPSVIS-----KATEKETFTDSNQSKTKKPGISDVTDYSDRQSDMDKDEATYSSQDH 56
 QY 524 WYPOSGSPPESTRNRSSTPAIRASDPATPIIPGCTTEISKGTGLSTY---- 579
 DB 57 QYPKQ-----ESSSVTSMKMFTESSSPK-----PGDJFEVELAKNDLSISYVLEFD 109
 QY 580 -GGSTFL--GAFTIHEVEEGACDGRMAGDOILEVNGIDLKATHEAIVNLRORP 636
 DB 110 KGGVNTSVHHGIIYKAVIPQGAESDGRHKGRVLAIVNGVSLGATHKQAVETLRNG 169
 QY 637 QYRVLTLXDEAPYKEEVY-----C-----DLETI 661
 DB 170 QYVHLLLEKGSPTSKENHVPVTPQCLSDQNAOGPEKVKTKTQKDYSPYTEENTFEV 229
 QY 662 ELQKKRPGKGLIS-----IVGRNDGVYSDIYKGIADPDGRILGDOILLVNGED 714
 DB 230 KLFKN-SSGLGFSFREDNLIPEQINASTYRYKKLFPGQPAESGKIDVDYLKNGAS 288
 QY 715 VNASQEAVALKCSLGTVTLEVGRIKAGPF-----HSERRPQT 755
 DB 289 LKGLSQOEYISALKGTAPVFLICRPPGVLPETIDTALLTLQSAQVLPNSSKSSQP 348
 QY 756 SOVSEGLSFTFPLSGSSTSSLESSKKNALASEIQLRTYEMKGGPTDSLGISAG 815
 DB 349 SCVEQ-----STSDENEMSDSKSKQC-----KSPSRDYSDSG 384
 QY 816 VQSPGLDVPITFAMHPTVANAQTKLRVGDHIVTICGSTEGMTHTQAVNLLKNSGT 875
 DB 385 SGE-----DDLVT-----APANISNTWSS 404
 QY 876 EMQVAGDVSIVTGHQBPASSLSFTGL-----TSTIFODDLR 917
 DB 405 ALHQTLNNVNSQAQSHHAKSQEDITCMFYYPQKIRPKPEREDSNPRLPDMAPGS 464
 QY 918 --PQCKS-----ITLERGPDG 931
 DB 465 YQPOSESASSSMKYHHIHSPTROENMTPLKNDLENHLEDFELEVELLTLIKSEK 524
 QY 932 -LGFSTVGGYSPHBDLPYVKTVPKGAASEGRKRDQIIIVNGSLGCTHEAIV 990
 DB 525 SLGFYTKG-----NORIGCYVHDVI-QDPAKSDGRILKPDRLKIVNDTIVNTHTDAVN 579
 QY 991 ILKRTGVTTLAV 1003
 DB 580 LRAASKTVRLVI 592

RESULT 14
 US-09-045-632-50
 Sequence 50, Application US/09045632
 Patent No. 6001575
 GENERAL INFORMATION:
 APPLICANT: Hugarli, Richard L.
 APPLICANT: Dong, Hualing
 TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
 TITLE OF INVENTION: GRIP-RELATED MOLECULES
 NUMBER OF SEQUENCES: 105
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
 STREET: 130 Water Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/045.632
 FILING DATE: 19-MAR-1998
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/041,016
 FILING DATE: 19-MAR-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Corless, Peter F.
 REGISTRATION NUMBER: 33,860
 REFERENCE/DOCKET NUMBER: 48147/1699-CIP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-523-3400
 TELEFAX: 617-523-6440
 INFORMATION FOR SEQ ID NO: 50:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1050 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-045-632-50

Query Match 6.6%; Score 334.5; DB 3; Length 1050;
 Best Local Similarity 21.9%; Pred. No. 4.2e-20;
 Matches 203; Conservative 129; Mismatches 321; Indels 275; Gaps 42;

QY 157 VELMREPSLSGISVGR-----GMGR-----LSNGEYMR--GIFKHYL 196
 DB 89 VELIKREGSLGHTIGGIDKDKKPRVSNLRREGLAARDLNVGYITSVNGIRLTRLR 148
 QY 197 EDSPPAKNCTL--KPGDRIVEAPSOSESPEKAPLCSVPPPSAPFAENGSDHTOSSASK 254
 DB 149 HBEII--TLKKNVERVY-----LEVEYE-----LPPAP-----ENNPRI 182
 QY 255 IQSDVDEDEFGYSKWNINERGTGLTGLHMTLEKGGHGLSLAG-----NKDRRMVY 310
 DB 183 ISKTYD-----VSLYKSGSFGVLKGAHEDLHKSRPLV 217
 QY 311 FYIGIDPNCAGADGRLOADELLEINGOILYGRSHQNA--STIKCAPSKVILFIKND 369
 DB 218 -LTYVAPGPANREGSLKGDRLSTXGIPRHASHATATATLQCC-----SHE 265
 QY 370 AVNQAVPCGNAVEPLPSNSENLOKNETEPTVTTSDAAYDLSSFKNVQHLLEPKDGGGLG 429
 DB 266 ALFOY-----EYDATPPTVANASGPLVEIAKTPGSGALGS 302
 QY 430 IAISEDTLSGYIKSLTEHGAATDGRKLVGDQIIIVADDEIVYGIPIKFTSLTAKM 489
 DB 303 LFTGSHRNKPATIDIRIKPASVVDKNGALHAGEHIIAIDGTSTENCSLVEATKLLASVTE 362
 QY 490 TYKLTTHAENPDSQ--AVPSAAGAASGKKNSQSILAWYPOSGSPPESTR 537
 DB 363 KVRLEI-LPAQOSRRLKPEAVRIQSRSEQLHMDP-CVPSCHSPRSHCRAPTAAPGGO 420
 QY 538 -NTSRSPATFAFSDPATPIIPGCTTEISKG--RT----- 572
 DB 421 DGRSVSTP--FSSPTNNPAP--CANASTLPRGMSRRTTAGRRORRKHRSLSLA 476
 QY 573 -----GLGLSTYGG--SDTLGATIIHEVEEGACAD 603
 DB 477 SSTVPGGOIVHTEVTEVVLGCDPLSGFGIQLGGIFATETLSSPLVFRTEPDSPAEK 536
 QY 604 GRMAGDOILEVNGIDLKATHEAIVNLRQ--PQVRVLTLYRDEAPYKEEVY--CDT 658
 DB 537 GLQVGDRLVLAINGIATEGCTHEANQLRDAALARKIYLEIFDVA-----ESVIPSST 592
 QY 659 LTELQKKRPGKGLST--VGRNDTGVFVSDIVKGIADPDGRILGDOILLVNGEDVR 716

```

Db 593 FHVKLPKRGEVLEGTITSSASRKRGEPLIISDIKKSVAHRTGLEPDKLAIIDNIRLD 652
QY 717 NASOE-AVAALLKCSLGTVTLEVGRIKAGPHSRERSQTS--QVSEGSLSFTFPLSGS 773
Db 653 HCPMEYAVOILPOCE-DLVKLKI-----RKDEDNDEDESSGAVS-----691
QY 774 STSELESSSKNNLASEIOGLRTVEMKK--GPTDSLGISIAGVGSPLGDPFIAMMH 831
Db 692 -----YVELKRYGCP--LGITIS-GTEEPFD--PIIISGLT 723
QY 832 PTGVAOQOKLRVGDRIYITIGTSTEGMTHQAVNLKNAGSITEMOVY-----A 881
Db 724 KRGLAERTGASMLGTAWPXSVSLSKGRPLSEAIHLQVAGETVTLKIKQIDRPLLRQ 783
QY 882 GADVSVVTHGHEBPASSLSFTGLTSTIFQDDIGPPOCKSITLERPGDLGFSIV-GGY 940
Db 784 SGSLSEASDVDEDPPEALKG--GLTTTHF-----SPAVSV--DSAVESMGSSATEGCF 833
QY 941 GSPHGDLPYVKTYFAKGAASEDRKLK 968
Db 834 GSGSGS---YTPQVAVRSVTPOEWRSSR 857

RESULT 15
US-09-045-632-49
: Sequence 49, Application US/09045632
: Patent No. 6001575
: GENERAL INFORMATION:
: APPLICANT: Hugent, Richard L.
: APPLICANT: Dong, Hualing
: TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
: NUMBER OF INVENTION: GRIP-RELATED MOLECULES
: NUMBER OF SEQUENCES: 105
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
: STREET: 130 Water Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/045,632
: FILING DATE: 19-MAR-1998
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/041,016
: FILING DATE: 19-MAR-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Corleese, Peter F.
: REGISTRATION NUMBER: 33,860
: REFERENCE/DOCKET NUMBER: 48147/1699-CIP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-523-3400
: TELEFAX: 617-523-6440
: INFORMATION FOR SEQ ID NO: 49:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1050 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-045-632-49

```

```

Query Match 6.5%; Score 331.5; DB 3; Length 1050;
Best Local Similarity 21.8%; Pred. No. 7.7e-20;
Matches 202; Conservative 130; Mismatches 321; Indels 275; Gaps 42;
QY 157 VELKREPSKSLGISIVGR-----GMGR---LSNGEYMR---GIFIKHVL 196

```

```

Db 89 VELIKREGSTGLTISGTEKDKGPRVSNLAPGGLAARSDLLNAGDYIRSVNGIRLRLR 148
QY 197 EDSAGKNGTL--KPGRIYAPASQSESEPEKAPLCSVPPPPSAFAEMGSDHTQSSASK 254
Db 149 HDEII-----TLKNVGERVY-----LEVEYE-----LPPAP-----ENPRI 182
QY 255 ISOVDKDEDEFGYSKMNIRERYGTLTGELHMELEKSHGSLAG-----NKDRSNV 310
Db 183 ISKTYD-----VSLYKBNSTGFLRGAGHDLKSRPLY 217
QY 311 FIVGIDPNGAAGKQRLQIADLELNGQLLYGRSHNA--SIIKCAPSKYKIIIFIRNKD 369
Db 218 -LTYVRPGGPANREGSLKVGDRLLSIXGIPLHGASHATAITLQOC-----SHE 265
QY 370 AVNMAVCPGNAVEPLPSNSNLQNKLEPEYVTSDAVDSLSPKNOHLELPDQGLG 429
Db 266 ALFOV-----EYDVATPDVANAAGPLVVEIATTPSALGIS 302
QY 430 IAISEEDTLGVIILKSLTEHGAATDRLKVGDOIILAVDEIVVGYPIEKFTSLKTAKM 489
Db 303 LTTGSHRNKRAITIDRIKPAVYDKNALHGEHIIAIDGSTHCSLVEATKILLASVTE 362
QY 490 TVKLTIAHENDSQ--AVPSAGAASGEKKNSSQSLMVPQSGSPESIR-----537
Db 363 KVRLEI-LPAOSRRPLKPEAVRIORSEQLHMDP-CVPSCHSPRSHCAPTMADGQ 420
QY 538 --NTRSRSTPAIFASDPATCIIGCETTIEISG-----RT-----572
Db 421 DQSRVSTP--FSSPTMANAP-CANASTLLPGRPSPRTTARRRORRKRHSLSLA 476
QY 573 -----GLGLSIVG--SDTLGAFIIEYERGAACKD 603
Db 477 SSTVPGQGIYHTEVTEVLCGDPGLGQIGFATETLSSPLVRFIEPSPARBC 536
QY 604 GRIMAGQIIEVNGIDLRKATHDAIVLROT--PQVRLLTYDEAPYKEEV---CDT 658
Db 537 GLLOVGDVLAINGIATEDGTMEQANOLLRPAALARKIVLEIEFDVA---ESVYPSGT 592
QY 659 LTELKPKPGKGLST--VGRNDTGFEVSDIYKGIADPDGRLIGDOIILVNGEDVR 716
Db 593 FHVKLPKRGEVLEGTITSSASRKRGEPLIISDIKKSVAHRTGLEPDKLAIIDNIRLD 652
QY 717 NASOE-AVAALLKCSLGTVTLEVGRIKAGPHSRERSQTS--QVSEGSLSFTFPLSGS 773
Db 653 HCPMEYAVOILPOCE-DLVKLKI-----RKDEDNDEDESSGAVS-----691
QY 774 STSELESSSKNNLASEIOGLRTVEMKK--GPTDSLGISIAGVGSPLGDPFIAMMH 831
Db 692 -----YVELKRYGCP--LGITIS-GTEEPFD--PIIISGLT 723
QY 832 PTGVAOQOKLRVGDRIYITIGTSTEGMTHQAVNLKNAGSITEMOVY-----A 881
Db 724 KRGLAERTGASMLGTAWPXSVSLSKGRPLSEAIHLQVAGETVTLKIKQIDRPLLRQ 783
QY 882 GADVSVVTHGHEBPASSLSFTGLTSTIFQDDIGPPOCKSITLERPGDLGFSIV-GGY 940
Db 784 SGSLSEASDVDEDPPEALKG--GLTTTHF-----SPAVSV--DSAVESMGSSATEGCF 833
QY 941 GSPHGDLPYVKTYFAKGAASEDRKLK 968
Db 834 GSGSGS---YTPQVAVRSVTPOEWRSSR 857

```

Search completed: July 12, 2001, 14:41:11
Job time: 214 sec

Fri Jul 13 15:00:11 2001

us-09-502-698-2.ra1

Page 13

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2001, 18:54:42 ; Search time 6307.92 Seconds

(without alignments)
11966.324 Million cell updates/sec

Title: US-09-502-698-3

Perfect score: 4880

Sequence: 1 cccggggccggggcagctg.....tctctcccccacccaact 4880

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rnd:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rnd:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_v1:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_v11:*
59: gb_v12:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_r01:*
95: gb_r02:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4185.6	85.8	6582	88 AF093419	AF093419 Homo sapi
2	2816.2	57.7	7516	95 RNMUPP1	AJ001320 Rattus no
3	2718.2	55.7	6288	94 MMU131869	AJ131869 Mus muscu
4	1745.6	35.7	1768	93 HSMUPP1	AJ001319 Homo sapi
5	1433.4	29.4	1681	9 AX011714	AX011714 Sequence
6	1213.2	24.9	2703	94 MMAR000168	AF000168 Mus muscu
7	614.4	12.6	169660	90 AL161449	AL161449 Human DNA
8	614.4	12.6	196607	79 AL353639	AL353639 Homo sapi

[illegible]

ALIGNMENTS

	AB044807	Homo sapiens
	AF060539	Mus musculus
G25381	human STS E	
G27450	human STS S	
AC013514	Homo sapiens	
AC008919	Homo sapiens	
AC013514	Homo sapiens	
AJ224747	Homo sapiens	
AJ001306	Homo sapiens	
AJ224748	Homo sapiens	
AC009919	Homo sapiens	
AC008455	Ceenorhabd.	
Z46792	Ceenorhabd.	
Z46792	Ceenorhabd.	
AF173450	Drosophila	
AF193199	Drosophila	
AF193199	Drosophila	
AC013734	Drosophila	
AC003537	Drosophila	
AE003472	Drosophila	
AL593802	Human DNA	
AL590374	Homo sapiens	
AL136458	Homo sapiens	
AL162739	Homo sapiens	
GC090170	Homo sapiens	
G20352	human STS A	
AL445143	Human DNA	
AC190787	Homo sapiens	
AC044779	Homo sapiens	
SF338650	Homo sapiens	
AF175292	Mus musculus	
AC044779	Homo sapiens	
U327476	Human channel	
AF166411	Rattus norvegicus	
AF233323	Homo sapiens	
X79676	H. sapiens P	
D21211	Human mRNA	
AB0289	H. sapiens P	

```

RESULT      1
AF093419
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE
gene
CDS

AF093419      6582 bp      mRNA      PRI      30-SEP-1998
Homo sapiens multi PDZ domain protein MUPP1 (MUPP1) mRNA, complete
cgs.
AF093419
AF093419.1      GI:3668409

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 6582)
Eng, L., Kravitsky, G. and Clapham, D.E.
Human homolog of MUPP1 protein
unpublished
2 (bases 1 to 6582)
Eng, L., Kravitsky, G. and Clapham, D.E.
Direct Submission
Submitted (21-SEP-1998) Cardiology, Children's Hospital, 320
Longwood Ave, Boston, MA 02115, USA
Location/Qualifiers
1..6582
/organism="Homo sapiens"
/db_xref="taxon:9606"
/issue_type="Brain"
/dev_stage="fetus"
1..6582
/gene="MUPP1"
47..6175
/gene="MUPP1"

```

[illegible]

OY 572 ttgatccagcaagcacttgatataatcttcttgctgctggcgccatctgtgaa 631
|||||
Db 2208 TTGATCCAGCAAGCACTGTGATTAAACGTTCTTTGGTGGCTGGCGCATGTGTA 2267
OY 632 agaatgagacacttctctggtgagccactcatgttctgaagcatgttcaacttgaa 691
|||||
Db 2268 AGGATGGAAGCACTTCTCTGGTGACCGACATCTTTGTAAGCATGTTAACCTTGGAAA 2327
OY 652 aacagcagctctgaggaagctgtagaagcactgaaggaagcacgcgtcaaggagctgtgaa 751
|||||
Db 2328 ACAGCAGCTTGAGCAAGCACTGTGAAGCACTGAAGGAGACCGTCAGAGCATGTGAGAA 2387
OY 752 taagtgctgctaagccttacccttaccaggaagaagtatgttctgttaaggagg 811
|||||
Db 2388 TAGGGTGTCTAAGCTTACCCCTTTACCCAGGAAGGTTATGTTCTGTGTAAGGAGG 2447
OY 812 attccttcttaaccacacacccctgtgaaggaagcaggcgtgtgacaacccctct 871
|||||
Db 2448 ATTCTTCTCTACCCACCACTCTCTGTGAAGGAAGCAGCGCTGGCTGACAAACCCCTCT 2507
OY 872 tcaaggcctgacttggcctgtgtggcacaatgactgacttagtagatgaaatccacat 931
|||||
Db 2508 TCAGGGCTGACTTGGCTGTGGGCAACAATGATGCTGACTTATGATGATGATGAT 2567
OY 932 ttgagcttccatctactctctgtgaagaatgacatctactcttaagcctctattat 991
|||||
Db 2568 TTGATCTCCATCTCTCTCTGTAAGATGACAGATCTACTCTACTCTAAGCTTATTTAT 2627
OY 992 ctctctacagcagctctctgtgtgactgagcctgaactatgcttccctccatctc 1051
|||||
Db 2628 CTCTTCAATGAGGAGCTTGTGTGTGATGGCTGATGATGATGATGATGATGATGAT 2687
OY 1052 ctctcaagagatgtatgtgaaatctctgtgactcaagctactgactgcaatgtctcgg 1111
|||||
Db 2688 CTCTTCAAGGAGCTTGTGTGTAAGATGATGATGATGATGATGATGATGATGAT 2747
OY 1112 aggaactatataccagaaatctctggaaagacagaaatgaaatgaaatgaaatgaa 1171
|||||
Db 2748 AGGAACCTATATACCAAGATCTCTGCAAGAGAGATGAGATGATGATGATGATGAT 2807
OY 1172 taagatgagggcctgtctgtgcttactataatgactacacacacacacacacacac 1231
|||||
Db 2808 TAAATATGGGGCTCTCTGTGCTTACTATTAATGATGATGATGATGATGATGATGAT 2867
OY 1232 aacaaacaatgaaatgtgaaacacaaatagtgagctgaaatcctaacttaagtgaa 1291
|||||
Db 2868 AACCAATATGAAATGTGAAACACAAATGATGATGATGATGATGATGATGATGAT 2927
OY 1292 ttaatacaagtcagaactctctgtgtgctaccgaaatcagctgaaagggcctgagc 1351
|||||
Db 2928 TTATATCAAGTGCAGAACTCTCTGTGTGCAACCGATTCACCTGGAAGGGCTGTGAT 2987
OY 1352 aacctgtgaaacagagctccctggcgtctgaatgctgagtgatgcaatgaaatgaa 1411
|||||
Db 2988 ACCGTCTGGAACAGAGCTCCCTGGCTGTAAATGCTGAGTGTGCTGCTTCAAAATGTAT 3047
OY 1412 ctaagaatcttctgaaagactatataatagcaaaagcaatctctagctagaatgaa 1471
|||||
Db 3048 CTAAAGAACTCTTTTGAAGGACTATTAAATATGCAAAAGGCAATCTTACCCAGGAATCA 3107
OY 1472 caagtgatgctaaataaagctgtggggaatgatactgtgaaagcaatcttaagtgag 1531
|||||
Db 3108 CAGTTAAGTCTAATAAAGATGCTTGGGATGATGATGATGATGATGATGATGATGAT 3167
OY 1532 ccatgagtcgaagatggcggaatgtgcaatgtgagactgcaatctgtccatgaaatgaa 1591
|||||
Db 3168 CCATTAAGTGCAGATGGCGGATTTGCCATTTGGGAGCTGCATCTTGTCCATTAATGAAGAT 3227
OY 1592 ctacaatagtgtaacaaatgcccgaagcagagactatgttgaagaagacatctctcaatg 1651
|||||
Db 3228 CTACCATGATGTTAAGCAATGCGCAGGCGAGGCTATGTTGGAAGACATCTCTCATATG 3287

OY 1652 gccctgacataaaatctactatgtgcctgcagaaacatttggaagagttcaaaatlaagc 1711
|||||
Db 3288 GCCCTGACATATAAATTTACTTATGTGCTGCGAAGCAATTTGGAAAGATTCAAATTAAGCT 3347
OY 1712 tgggaacaacatctggaagatgaaatgagcactgaaatcttctctacacactggaag 1771
|||||
Db 3348 TGGGACAACAATCTGGAAGATGATGCGACTGATATATTTTCTTCATTAACATGCGCAGAG 3407
OY 1772 acatccagaatctaccagaagcagaagagaagagagtggaagaagcgaactcaaaaca 1831
|||||
Db 3408 ACATTCAGAAATTAACCAAGCGAGGAAGGAGAGGCTGGAAGAAAGCGCAATCTCAAAACA 3467
OY 1832 cagcatatagcaatgtgaatcagcccaagcagggtggaactcgtggaagagaacaaacaat 1891
|||||
Db 3468 CAGCATATGCAATTTGAAATGACCCAGCGGGGTGAACTGTGGAAGAAACCAAGCAAT 3527
OY 1892 ccttagacatcagactgtgtgagcagagatgaggagctgctgaaagcaatgagaag 1951
|||||
Db 3528 CCTTAGGATCAGCATTTGTGTGTGAGCAGAGGATGGGAGTGGCTTAAGCATGAGAGAG 3587
OY 1952 tgaagaagggaatcttcatcaaaacatgtctggaagatagctcagctggcaaaatgaa 2011
|||||
Db 3588 TGATGAGGGGCAATTTTCATCAAAACATGTTCTGGAAGATGATGATGATGATGATGAT 3647
OY 2012 ccttgaacactggaagatgaaatcgtgag 2038
|||||
Db 3648 CTTTGAACCTTGAGATGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 3707
OY 2039 2038
Db 3708 ATGAACAAGCTGTGGAAGCATTTGGAAGAGCAGCAACCTGTAGCTTATGTATGATGAGA 3767
OY 2039 2038
Db 3768 GCATTATAAACAGACCAAGAAATCCCTTGGCTTGGCTGCAACACTTTACCTTA 3827
OY 2039 2038
Db 3828 AGTACAACTTGACGACACTAACCCATTTGCTGACTCTCTACAAATCAAGCGCGACMAG 3887
OY 2042 caccagtcagtcagagtcagagccagagaagagtcacatctgtgcaagtggtgccccccc 2101
|||||
Db 3888 CACCAGTCAGTCAGAGTCAGAGCCAGAGAAAGCTTCATTTGTGCAAGTGGCCCCACCCC 3947
OY 2102 ctccctcagccttgcggaatgagtgatgatacacaacagctcactgcaagcaaaatct 2161
|||||
Db 3948 CTCTTCAAGCTTGTCCGAAATGGGTAGATGATGATGATGATGATGATGATGATGAT 4007
OY 2162 cacaagaatgtgacaagaagagatgagttgtgtacagctggaagaaatatacagaagcgt 2221
|||||
Db 4008 CACAAGATGTGACAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 4067
OY 2222 atggaacccaaacagcagctgcatatgactgaaatggaagagagtcagtagtggtg 2281
|||||
Db 4068 ATGGAACCTTACAGAGCGAGCTGATATGATGATGATGATGATGATGATGATGATGAT 4127
OY 2282 gcttaagcttctgtggaacaaagccgatacagaatgagtgcttctcaatgagggagatg 2341
|||||
Db 4128 GCTTAAGTCTTGTGGAACAAAGACCGATCCAGATGATGATGATGATGATGATGATGAT 4187
OY 2342 atccaaatgagcctgcaagaaagaatggtcgaatgcaaatctcagaatgagctccagaga 2401
|||||
Db 4188 ATCCAAATGAGACTGACAGAAAGATGATGATGATGATGATGATGATGATGATGATGAT 4247
OY 2402 tcaatgltcagaatcttataatggaagaagtcacagaatgctcactcaatcaatlaaagt 2461
|||||
Db 4248 TCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4307
OY 2462 cccctctcaagtgaaataatcttatacagaataaagaatgcaatgataatgagtgagc 2521
|||||
Db 4308 CCCCTTCTTAAATGAAATTAATTTTATCTAGAAATTAAGATGATGATGATGATGATGAT 4367
OY 2522 tatgtccctggaatgagtagaaccttgccttcaactcagaanaatcttcaaaataag 2581

```

|||||
Db 4368 TATGCTCGGAATGAGATGAAACCTTGCTTAACTCAGAAATCTTCAAAATAAG 4427
OY 2582 agagaagggcgaagcttactactctctgtagagcttgtagcctaagtctatttaaatg 2641
Db 4428 AGACAGAGCCCAAGCTGTAAGTCTGATGAGCTGAGCCTGATTCATTTAAATG 4487
OY 2642 tgcacaactctggaagctcccaaggatcagggggtctgggtattgtcatacagcgaag 2701
Db 4488 TCGAACATCTGAGCTTCCCAAGATCAGGGGGTTTGGGTATTGCTATCAGCAAGAG 4547
OY 2702 atcaactcagtgtagatcaataaagagcttaacagagcattggttagcagcagatg 2761
Db 4548 ATACACTCAGTGAATCATCTTAAAGCTTAAACAGAGCAAGGGGTAGCACCGGATG 4607
OY 2762 gacgacccaagctcggaagatcagaactgagctgtagatgtagaattgtgttgtagc 2821
Db 4608 GAGCACTCAAAAGTGGAGATCAGTACTGCTGTAGATGATGAATTTGTTTGTTTACC 4667
OY 2822 ctatlgaaaagtattatagcctctgaaagacagcaaeagtagaataacttaacatcc 2881
Db 4668 CTATTGAAAAGTTTATTAGCTTCTGAAGACAGCAAGATGACAGTAAACTTACATCC 4727
OY 2882 atgctggaatcccaagattcccaagctgttcttcaagcagctggtgagccagtgagaaa 2941
Db 4728 ATGCTGGAATCCAGATTCAGAGCTGCTCTTCAAGACAGCTGAGTGCAGCCAGAGAAA 4787
OY 2942 aagaagacagctcccaagctctctgtagtggcccaagctgagctcccaagacagagatca 3001
Db 4788 AAAAGAACAGTCCCACTCTCTGATGCTCCACAGCTGCTGCTCCCAAGAACCGAGTCCA 4847
OY 3002 tccgaaatacaagacagatcaatacaacacagcaalttctgtcttcaactcgaactgccc 3061
Db 4848 TCCCAATACAGAGATCATCAACACAGCAATTTTGTCTGATCTCAGTCAACCTGCC 4907
OY 3062 ccatatccctggtctgcaaaacacatctgagatcttccaaaggcgaacagggctggcc 3121
Db 4908 CCATTATCCCTGGCTGGAAACAACATCGAGATTTCAAAGGGCGAACAAGGCTGGGCC 4967
OY 3122 tgaagatctgtaggagctcagacagctgctggtgtgcttattatcatgaagtattatg 3181
Db 4968 TGAGCATCTGTTGGGGGTTCAACACGCTGCTGGGTGCCATTATTATCATGAAATTTATG 5027
OY 3182 aagaagagcagcatgtaaaagatgaaagactctgagcttgaagatcagatcttagagtg 3241
Db 5028 AAGAAGAGCAGCATGTAAAGATGGAAGACTCTGGCTGAGATCAATCTTAGAGGTGA 5087
OY 3242 atggaatctgtagaagagccacatgtagaagcaatcaatgctctgagaagagcgc 3301
Db 5088 ATGGAATTTGACTTGAGAAAGGCCACACATGATGAAGCAATCAATGCTCTGAGACAGC 5147
OY 3302 caagagagctgagcctgacactctacagagatgtaggcccatacaaaagaggaagtg 3361
Db 5148 CACAGAGAGTGGCCTCAGACTCTACAGAGATGAGGCCCTACAAAGAGAGAGAAATGCT 5207
OY 3362 gtagacccctcaactatgagctgagagaagagccgggaaagagccttagatlaagatlg 3421
Db 5208 GTGACACCCCTCATTATGAGCTGCAGAAAGCGGGAAAAAGCCCTTAGATTAAGTATG 5267
OY 3422 ttggttaaaagaaagataacgaggtattgtgtltagacattgtlcaaaagaggaattg 3481
Db 5268 TTGGTAAAGAAAGATGATGAGATTTTGTGACATTTGTCAAAGGAGGAATTTACAG 5347
OY 3482 atcccgatggaagatgataccagagagagacagatatattgttgaatggaagagttc 3541
Db 5328 ATGCGCATGGAAGACTGATGACAGGAGACCAAGATATTAAATGGAATGGAGAACCTTC 5387
OY 3542 gtaatgctcccaagagcggttgccgcttgcttaagtgcttccatgagcagtaacct 3601
Db 5388 GTATATGCCACCAAGAGCGGTTGCCCTTTGCTAAAGTGTTCCTTAGGACAGTAACT 5447
OY 3602 tggaggttggaaagatcaaaagctggltccattccattcagagaagagagcaltcaaa 3661
|||||
|||||
Db 5448 TGAAGTTGGAAGATCAAG----- 5468
OY 3662 gccaggtgagtgaaagcagcctgtctcttccattccactctctggtatccagtaac 3721
Db 5469 -----CTGGATCCAGTACAT 5483
OY 3722 ctgagctactgaaagtagctcaaaagaaatgcatctggaactcgaatacaagagatcaa 3781
Db 5484 CTGAGTCACTGGAAGATAGCTCAAGAAAGAAATGCATTGGCATCTGAATAACAGGATTTAA 5543
OY 3782 gaacagctcgaaatgaaagggccctactgaactcaggatggaatacgaatcgttgagag 3841
Db 5544 GAACAGTCGAAATGAAAGGGCCCTACTGACTCAGTGGGATATGAGATCGCTGGAGAG 5603
OY 3842 tagcagcccaacttggtagatgtgcttattatctgaatgtagtaccaccaactgagttg 3901
Db 5604 TAGCACCCCACTGGTGTATGCTGCTATATTATTGCAATGATGACCAACTGGAAGTTG 5663
OY 3902 caagcagaaccacaaactcagagctgaggatagagattgtacacatctgtagacatcca 3961
Db 5664 CAGCAGAGACCCAAAACTCAGAGTTGGGATAGGATGTGCACATCTGTGCACATCCA 5723
OY 3962 ctgggggcatgagctcacaccacagtaaacctcagaaatgcatctggtccatctg 4021
Db 5724 CTGAGGCGATGACTCACACCCAAAGCAGTTAACTTACGTAAGAAATGCAATCTGGCTCAATG 5783
OY 4022 aaatgcaggtgtgtctggaagagacgtgagtggttcaagctcatcaatcagagcctg 4081
Db 5784 AAATGCAGGTGTGCTGTGAGAGACGTAAGTGTGTACAGTCACTCAGCAGAGACCTG 5843
OY 4082 caagctccagctcttcttcaactcagtgagctgagctgaacagatatactcagatgattag 4141
Db 5844 CAAGTCCAGCTCTTCTTCTACTGGCTGACGCTCAACAGATATTTACAGATGATTTAG 5903
OY 4142 gacccctcaatgtaagtattatctactatagagagagacagatgagcttgagcttagta 4201
Db 5904 GACCTCTCAATGTAACTTATTAATCACTAGAGCGAGACCGAGATGCTTGAAGCTTCA 5963
OY 4202 tagttgagagatagcagccctcatgtagagcttaacatctatgcttaaaacagtgctg 4261
Db 5964 TAGTTGAGAGATATGCGACGCCCTCATGAGACTTATGCCATTATGTTAAACAGTGTG 6023
OY 4262 caaaggagcagcctctgaagacgagcgtctgaagaggggagatcagatcattgctga 4321
Db 6024 CAAAGGAGAGCGCTCTGAAGACGAGCTGCAAAAGGGGAGATCAATCTTCTGTCA 6083
OY 4322 atgggacaggtctgaagaggtccacatgaagaagctgtgtgcatctcttaaaagagcaa 4381
Db 6084 ATGGGACAGACTTGAAGAGAGTCAACCATGAAGAGCTGTGGCCATCTTAAACGAGCAA 6143
OY 4382 aagcactgtcaacttgaatgtctcttgaatgtgctccagaaatgaaaccaaccacaa 4441
Db 6144 AAGGCACTGTCACTTTATGTTCTCTCTTGAATTTGGCTGCGCAATTTGAACCAACCAA 6203
OY 4442 cccctagctcaactctactgtlaaagaaatgcaatggtccctgacaaatttlaatgctg 4501
Db 6204 CCCCTAGCTCACTCTTACTGTAAGAGAAATGCACTGCTGTCGATCAATTTTATGCTG 6263
OY 4502 ttcaagccgggtctcaaaactgtatgggggaaataaacaactaagtctctctctcactc 4561
Db 6264 TTCAAGCCGGGCTTCAAAACTGTAGGGGGGAAATTAACACTTAAGTTTCTTCTCAATCT 6323
OY 4562 agaaagcttctccttactgcaaacctacatacttctctctctctctctgcatcttg 4621
Db 6324 AGAAATCTTTCTTCTTACTGCAACCTTAACATATTCTTTCTTCTTCTTCTTCTTCTGA 6383
OY 4622 acttaaaagaaagaaatattgttgaagtgaaatcgttcttcttcttctgtagagatata 4681
Db 6384 ACTTAAAGAAAGAAATATTGTGTAGTGAATCTGTTTATTTTGTGAGATATCTAA 6443
OY 4682 tgtttgttagtcaatgaggaagatataacatgctaaagctggttagtataaagaaaga 4741
Db 6444 TGTTTGTAGTCAATGGGCAAGAAATTTATACATGTAAAGCTGTGTTAGTTAAAGAAAGA 6503

```


[illegible]

Db 3057 CCCAGAGAGCTTTTGAAGGCGCCGTTACTATAGCAAAAGGCGCTCCAGCCTAGTATGA 3116
Oy 1472 cagtagtgcataaagaatgagcttgaggatgatacgttcgaagcatatcaatgagtg 1531
Db 3117 CAGTAAGTCCCAATTAATGATGCTGCGGAGTGTGCGGAGCATTCATTCATGAGGGGCG 3176
Oy 1532 ccattagtcgaatgagcagatgagcttgagcttgagctgcatctgtccatlaaagaagc 1591
Db 3177 CCATTAGTGTGATGCGCGGATGCTGTGGGCGACTGCTTTTGTCCATTATGAGAGAT 3236
Oy 1592 ctaccatagtgtaaccgaatgcccagcagcagcagcagcagcagcagcagcagcagcagc 1651
Db 3237 CCACCATCAGTCTACCAATGCGGAGGCGGCGCATGCTGAGAGACATTCCTAATTTG 3296
Oy 1652 gccctgacaataaattactatgctgtcctgcagaacattggaagagtcgaataagc 1711
Db 3297 GACCTGATATCAAAATTTACTTACGTGCTGCGAGAACATTTGGAGAGTTCAGAGTAAAT 3356
Oy 1712 tggagcaaatctggaagagtaatgagcagcagcagcagcagcagcagcagcagcagcagc 1771
Db 3357 TTGGTCAACAGCTTGAGAGATATGCGACTGATATTTTCTTCATACACTGGCAGAG 3416
Oy 1772 acattccgaattaccagagcag 1831
Db 3417 ACATTCCAGAACTCCAGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3476
Oy 1832 cagcatatagcaattggaatcagccagcagcagcagcagcagcagcagcagcagcagc 1891
Db 3477 CTGGCTTACAGCGAGTGCAGCCAGCCAGCGAGGTGGAATTTGGAGAGAACCAACAAAT 3536
Oy 1892 ccttgagatcagcagctgtgtgagagagagagagagagagagagagagagagagagag 1951
Db 3537 CTTTGGGCAATCGAGCTTGTGTGTCAGAGATGGGAGCCGCTGACCAATGAGTGAGG 3596
Oy 1952 tgaatagagagcatttcac 2011
Db 3597 TGATAGGGGCTTTTTCATCAACAACATGTTCTTGAGAGACAGTCCAGCTGCAAAAATGGA 3656
Oy 2012 ccttgaacactggaatagatcagtag----- 2038
Db 3657 CTCTGAGCCAGAGACAGAAATTAATTGAGGTGATGAGTGCAGCTCAGAGATGCAAGCC 3716
Oy 2039 ----- 2038
Db 3717 ATGAACAAGCTGTGGAAGCCATTGCGAAAGAGCAACCTGTAGTATTATGTACAGA 3776
Oy 2039 ----- 2038
Db 3777 GCATTATAACAGCAAGAAATCCCTTTGCCCTTCTTGGCGGACAGCCCTTTACCCCTA 3836
Oy 2039 ----- 2038
Db 3837 AGTACAGCTTACGACACTAATCCATTGACAGCTCTCTCCAGCTCAGCACTGACAGCAAG 3896
Oy 2042 caccagatcagtcagagtcagagtcagagtcagagtcagagtcagagtcagagtcagagtc 2101
Db 3897 CACCCAGCCAGTCAGATCTGAGACAGAGAACCTGCAATGTGCAATGTTCTCCCTCTCT 3956
Oy 2102 ctccctcagccttgcgaaacagcagcagcagcagcagcagcagcagcagcagcagcagc 2161
Db 3957 CTCCTTCAAGTGTTCAGAAATGGGATGATGTTGACAGCAGCAGCTGCAACCGCAGTCT 4016
Oy 2162 cacaagaatgagcaaaag 2221
Db 4017 CAGAGAGTGAAGCAAAAGAGATGAGTGTGTTACAGCTGGAAGAAATATCCAGAGCGCT 4076
Oy 2222 atggaacacttaacagagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2281
Db 4077 ATGGAAGCTGACAGGTGCTGCTCCATGATGAGCTGAGAGAAAGGCGAGCAGCGGTTGG 4136
Oy 2282 gctaagcttctggaacaaagacagcagcagcagcagcagcagcagcagcagcagcagc 2341

Db 4137 GTCTAGTCTTGTGGGAACAGACCCAGCAGATGATGTTATAGTGGGATTTG 4196
Oy 2342 atccaaatgagcctgcaagaaagaatgctgcatatgcaatgcaatgcaatgcaatgcaatg 2401
Db 4197 ATCCCACTGAGACAGCGGAGAGATGGCCGACTCAGATTTGCCAGCAGCTTTAAGAGA 4256
Oy 2402 tcaatgagcagatttataatgagaagaatcagagagagagagagagagagagagagagag 2461
Db 4257 TCAATGGCCAAATTTGATGAGCAGAGATGATGATGATGATGATGATGATGATGATGAT 4316
Oy 2462 cccctcctaagtgaaataatttatacagaataaagaatgcaatgcaatgcaatgcaatg 2521
Db 4317 CTCCTATCAAAATTAATAATTTTATCAGAAATGACAGATGACATGACATGACATGAC 4376
Oy 2522 tatgtccctggaatgagtagaagccttgcccttcaactcagaataatcgaataaag 2581
Db 4377 TATGTCTGGGATTCAGAGAGCTCTCCGCTTCACCTCAGACAGTCTCAGAACAG 4436
Oy 2582 agacagagcgaactgttaactctgagcagcagcagcagcagcagcagcagcagcagcagc 2641
Db 4437 AGGTGGAACATGACAGTACTGATGCTGCTGCGGAGCTCAGCTCAGCTCAGCTCAG 4496
Oy 2642 tgaacatctgagccttcccaagagtcagggggttggtgatactgcatcagcagaag 2701
Db 4497 TGTACACAGTGAAGCTTCTTAAGATCAAGAGAGTTAGGCTTATGCTATCTGTGAGAG 4556
Oy 2702 ataacctcagtgagatcacaataaagagccttaacagagagagagagagagagagag 2761
Db 4557 ACACATCAACCGAGTACATGATGGAAGCTTAACGAGCAGCGGAGAGCAGCAAGAG 4616
Oy 2762 gacagcctaagtcgagatcagcagcagcagcagcagcagcagcagcagcagcagc 2821
Db 4617 GAGGCTCAACCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4676
Oy 2822 ctattgaagaatttataagccttgaagaacagcagaatgacagtaaacattacatcc 2881
Db 4677 CTGTTGAAGAATTCAATGAGCTTCTGAAGAGCGCAAGCACTGTAATGAGCTGTTCC 4736
Oy 2882 atgtcagagatccagatcccaagcctgctcctcagcagcagcagcagcagcagcagc 2941
Db 4737 GAGCTGAGATGACAGCTTGTCCAGCTGTTCTTCACTGATTAACAGTCACTGAGAGAA 4796
Oy 2942 aagaagacagcctccagcagcagcagcagcagcagcagcagcagcagcagcagc 3001
Db 4797 GGAAAGACACCTCCAGACTCTGCAAGTCCA-----CTCCAGACTGGAACCA 4847
Oy 3002 tccgaatacagaagcagatcacaacacacacacacacacacacacacacacacacacac 3061
Db 4848 TCCCAAGTCCAGAGCTGCTCCACACAGAGTGTGCTTGTGCTGACCTGACCTGCC 4907
Oy 3062 ccattatccctgctgagaaac 3121
Db 4908 CCATATCCAGAGCTGTGAGACAAATTTGAATTTCCAAAGCCAAAGAGCTGGGAC 4967
Oy 3122 tgaacatgctgagagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3181
Db 4968 TGAATTTTGTGGGCTGACACACACTGCTGGGTCTTATTTATCCATGAGATTATG 5027
Oy 3182 aagaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3241
Db 5028 AAGAGGACAGCAGTGAAGATGAGACTATGGCTGAGAGACAGATTTAAGAGTAA 5087
Oy 3242 atggaatgacttgagagagcagcagcagcagcagcagcagcagcagcagcagcagc 3301
Db 5088 ATGGATTGACTTGGAAAGGCTTACACATGATGAAGCAATTCATTCAGAGAGAGCC 5147
Oy 3302 cagagagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3361
Db 5148 CTCMAAGAGTACGAGCTGCTTACCGAGATGAGCCCATTAACAAGAGAGATGCT 5207
Oy 3362 gtgaacacctacattgagctgcagaagagcagcagcagcagcagcagcagcagcagc 3421
Db 5208 GTGATACCTTCAACATGAGCTGAGAGAGCCAGGCAAGAGCTTGGTGTGATATG 5267


```

Oy 3412 ttaagatctgttgcgaagaacagatctgagatcttctgacacatgctcaaga 3471
Db 361 TTAAGTATTTGTTGTTAAAGAAACATACGTGAGATTTGTTGACATTTTCAAGA 420
Oy 3472 ggaattgacagatcccgatggaagactgacccagagacagatatcttggtaag 3531
Db 421 GGAATTCAGATCCGATGAGAGACTGATGACAGGAGACAGATATTAATGATGAG 480
Oy 3532 gaagacgttctgtaagctcccaagaagcggttgcgcttctgtaagcttccagg 3591
Db 481 GAAAGCGTTGTAATGACACCCAAAGCGGTGCGCTTTCCTAAAGTTCCTTAGGC 540
Oy 3592 aacgaaccttggagatggaagaaatcaaacgtggtccatctccatccagagag 3651
Db 541 AACGTAACCTTGGAAGTTGGAAGATCAAGCTGGTCCATTCATTCACAGAGAGCCA 600
Oy 3652 tctcaaacacagcaggtgagtgaaagcgacctgtcttcttccatcttccatctg 3711
Db 601 TCTCAAGCAGCCAGGTGAGTGAAGCGACCGCTGTCATCTTTCATCTTCCATCTG 660
Oy 3712 tccagtaacatgagatcactggaagtagctcaagaagaatgcatctgcaata 3771
Db 661 TCCACTACATCTGAGTCACTGGAAGTACTCAAAAGAAATGCAATGGCATCTGAATA 720
Oy 3772 cagggattaagacagctgaaatgaaagagcgccactgactgactgaggaatcag 3831
Db 721 CAGGATTTAAGAACAGTGCATAATGAAGAGCGCTTACTGACTCATCTGGGAATCAG 780
Oy 3832 gctggagagatgagcagccactgtgtgacatatacttctgcaatgacagca 3891
Db 781 GCTGAGAGGTGAGCGACCGCACTGTGTGATGCTATATTTATGCAATGATGACCCA 840
Oy 3892 actgagatgacagcagacagacccaaactcagatgtgggagataagatgcatcgt 3951
Db 841 ACTGAGGTGACGACACAGACCCAAACACTCAAGTGGGATGAGATGCTCCACATCTGT 900
Oy 3952 ggcacatcactgagagcgatgacacaccccaagcagttacacttcaaaatcatc 4011
Db 901 GGCACATCTCAGAGGCGATGATGATCACACCCAGCAAGTTAACTTGTGAAAATGCA 960
Oy 4012 ggtccatgaaatgacaggtgtgtgctggaagagacgtgagtgctgacagtgat 4071
Db 961 GGTCTCATTTGAATGACAGGTGCTGTGAGAGAGAGTGAAGTGTGTCACAGTATCAT 1020
Oy 4072 cagagcgctgacagctcactgcttcttcttcaatgagcgctgacacagatctt 4131
Db 1021 CAGGAGCGCTGCAAGTTCACAGTCTTCTTCTTCACTGGGCTGACGTCAAGCA 1080
Oy 4132 gatgattagagccctcccaatgtaagctatatacactagagcagagacagatgct 4191
Db 1081 GATGATTTAGAGCCTCCCTCAATGTAAGTCTATTACACTAGACGAGACCAATGGCTTA 1140
Oy 4192 ggtctcagatagttgagagatagcagccctcactgagagacttaaccaattagt 4251
Db 1141 GGTCTCATTTGATGTTGAGGATATGAGCAGCCCTCATGAGACTTACCCATTATG 1200
Oy 4252 acaaggttgcgaagagcgagccctcgaagcgagcgtctgaagaagcgatcagat 4311
Db 1201 ACAGTGTGTTGCAAGAGGACAGCCTCTGTAAGAGCGAGCTGTAAGAGGGGCGAT 1260
Oy 4312 attgtctcaatgagcagatctagaagagatcccaatgaaagagcgtgtccatcct 4371
Db 1261 ATTGCTGCAATAGGCGACAGTGTAAAGAGATCACCCATGAAGAGCTGTTCCATCT 1320
Oy 4372 aaacggacaaagcagctgacatctgtgtcttcttcaatgctgtgctcagaatga 4431
Db 1321 AAACGGAACAAAGGACAGTCACTTGTGATGCTTCTCTTGAATGGCGCAGAAATGA 1380
Oy 4432 accaaccacacccctagctcactctactctgtaagaagaatgcatgctgtccgaat 4491
Db 1381 ACCAACCCAAACCCCTAGCTCAGCTGTACTGTAAGAGAGATGCACTGGTCCGCAAT 1440
Oy 4492 ttaatctgttctcagcggtcttcaaaactgtaaggggagaaacttaagttctt 4551

```

```

Db 1441 TTAATCTGTTACACCGCGCTTCAAAAGTATAGGGGGAAATAACTTAAGTTCT 1500
Oy 4552 ttctcactagaaatgcttcccttactgacacacacatcaatcttcttctctg 4611
Db 1501 TTTCTCATCTAAGAAATGCTTCTTCTTACTGACAACTCAATCATCATTTTCTTCT 1560
Oy 4612 catlttgaaacttaagaagaagatatctgtgagtgaaatccgcttcttcttctg 4671
Db 1561 CATTTTGAACCTTAAGAGGAGAAATTTGTGAGGGAATTCGTTTATTTGTGG 1620
Oy 4672 agatctcaatcttctgtaagcagatggaagaaatcttaacatgacagctgtgag 4731
Db 1621 AGATTTCAATTTTGTGATGATACATGGCAAGATTTATACATGCTAAGCTGATGA 1680
Oy 4732 taaagaagataatcttaagagcacaagaagaatgctcagtaagtgagatga 4791
Db 1681 TAAAGAAAGATTAATCTTAAGAGTAAACCAAGAAATGCTTCAATGATGATGA 1740
Oy 4792 atgaaatataaataaagaagaaa 4817
Db 1741 ATGAAATATATAAAAAAAAAAAAAA 1766

```

```

RESULT 5
AX011714 1681 bp DNA PAT 06-SEP-2000
LOCUS AX011714 Sequence 112 from Patent WO955858.
DEFINITION AX011714
ACCESSION AX011714
VERSION AX011714.1 GI:9998238
KEYWORDS
SOURCE
ORGANISM human.

```

```

REFERENCE Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
Pillarsky,C.
TITLE Human nucleic acid sequences obtained from pancreas tumor tissue
JOURNAL Patent: WO 955858-A-112 04-NOV-1995;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METZGEN GES FUER GENOMFORSCHUN
(DE); PILLARSKY CHRISTIAN (DE)
FEATURES
LOCATION/Qualifiers
source 1..1681
/organism="Homo sapiens"
BASE COUNT 498 a 330 c 445 g 408 t
ORIGIN

```

```

Query Match 29.4%; Score 1433.4; DB 9; Length 1681;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1462; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

```

```

Oy 2106 ttcaagccttgcgaagaatgagtgatgacacagaatcctggaagaatctcaca 2165
Db 1 TTCACCTTTTGCAGAAATGAGTATGATCACACAGTCAATCTGCAAGCAAAATCTCACA 60
Oy 2166 agatgagcaagaagatgagttggttacaagcgtggaataatcagaagcgtttag 2225
Db 61 AGATGTGCAAGAAAGAGATGAGTTGGTTACAGCTGGAATAATATACAGACGTTATAG 120
Oy 2226 aacctaacagcgagctgacatagatgtaactggaagaagtcatagtgttggcct 2285
Db 121 AACCTTAACAGCGAGCTGACATATGATTAATGAACTGAGAAAGGTCAATAGGTTGGCCT 180
Oy 2286 aagcttgcggaagaagaacagatccaagatgagtggtctcacaagtgaggatgac 2345
Db 181 AAGTCTTGCGGAACAAAGACCGATCCAGAGTGAAGTGTCTTCATAGTGGGATTGATCC 240
Oy 2346 aatgagcctgaggaagaagatgctgcatgtaactgacagatgagctctagagatcaa 2405

```


Db 121 AACTCCGACACCTCCGATCCCA-----GCTCCAGACTGGAAACCCATCCCAACT 171
 Qy 3010 acaagacatcatcaacaccagaacttttcttctgacccgcaactgccccattac 3069
 Db 172 ACAAGAGCTCCCTCCACACACAGCTCTTGGTTCTGACCTGCCATGCCCATCATT 231
 Qy 3070 cctgctgtaaaacacacccctcgagatttccaaaggcgaaacgggtggtcctgagc 3129
 Db 232 CCAGCGCTGGAACAACAATTGGATTTCGAAAGGCCAAAGCGCTGGAGCTGATATT 291
 Qy 3130 gttgaggtctcagacacagctctggtggtccttatccatgaagttatgaaga 3189
 Db 292 GTTGGGGCTCAGACACACTGCTGGGTCTATTATTATCATGAATTTATGAAGGGA 351
 Qy 3190 gaagcatgtaaagatggaagactctggtcgtgagatcagacttgaagtgaaat 3249
 Db 352 GCAGCGCTGTAAGATGGAAGACTATGGCTGAGACCAATTTTAAAGGTAAATGGATT 411
 Qy 3250 gaattgaggaaggccacatgatgagaatcaatgatcttgaagacgccaagaga 3309
 Db 412 GACTTGGCAAGAGCTACATGATGAAAGCAATGATCTGAGGAGAGCGCTCAAGA 471
 Qy 3310 gtgcgctctacactctacagagatgagggcccaatacaagagaggaagtgatgac 3369
 Db 472 GTACGGCTGACGCTCTACAGAGATGAGGCCCATACAAAGAGAGATGCTGTGATAC 531
 Qy 3370 ctaactatg-----agctcgaagaagacgggaagagcctgaagttaaglatgt 3423
 Db 532 TTCCACCATGAGACTGAGCTGAGAGAGAGCGCAAGGCGCTTGGTTGATTTGTT 591
 Qy 3424 ggttaaaagaagatactgagatgtgtgtcgaagatctgcaaaaggaatgtgaat 3483
 Db 592 GCGAANAAGAAATGACACTGTGAGTATTGTATCAGACATTGTCAAGGAGGCAATTGACAG 651
 Qy 3484 ccgcatgtgagaactgtatccagaggaacagatattatgtgtaagtgggaagctgtc 3543
 Db 652 GCCGATGGAGACTGATGCAAGGGGACAGATTTTAAATGATGAGAGATGTCCTG 711
 Qy 3544 aatgctcccaagaagaggtgtgcgtctgtcgaagtgttccctgaagcagtaactgt 3603
 Db 712 CATGCCACACAGAGCACTGCTGCTGCCCTGCTAAAGTGTCTTCCATGATGATGAC 771
 Qy 3604 gaatttgaagaatacaagctgtgcattccattcagagagagggcactcacaacag 3663
 Db 772 GAGGTTGGAAAGTCAAACTGCCCATTCACACTCAGAACGGAGGCTTCTCAAAACAG 831
 Qy 3664 caggtgagtgaaagcagcctgtctcttcttcaacttctcactctctgagtcagta 3723
 Db 832 CAGGTGAGTGAAGAGCAGCTGTCTTACATCTTACATCCCTCCACTTCTGGAATTAATCATCA 891
 Qy 3724 gagtcactgtgaaagtgtcctcaaaagaatgtcgtcgtcgtgaatcaaggaatga 3783
 Db 892 GAGTCTTTGAAATGAACTCGAAGAAATGCAATTCGAAATTCGAAATTCGAAATTCG 951
 Qy 3784 acaagtgaatgtaaaaagggccctactgactcaactggaatcagcctgctgagga 3843
 Db 952 ACAGTGAATTAATAAAGGGGCTGCTGACTGCTGGACTCGACTGACATTCGTTGAGAGAG 1011
 Qy 3844 ggcagccactgtgtgtgtgtcctatattatgtcaatgtcgaacccaactgtgagttga 3903
 Db 1012 GCGAGCGCTCGAGATGCTCTATATTATTATTCGATGATGACCAATGATGTTGGA 1071
 Qy 3904 gcaagaccccaaaactcagagttggtgagatagatgtcacaactcgtgtgacacact 3963
 Db 1072 GCTCAAAACCAAAACTCAGAGTGGGATGAGATTGACCAATTCGTGGCAATTCAG 1131
 Qy 3964 gaggagcatctacacaccccaagagtaactactaataaagtcatctggtcctcatgga 4023
 Db 1132 GATGGATACCTACACACAGCGGTTAACTTATATATAAATCCCTCAGGCTCATTTGA 1191
 Qy 4024 atgcaggtgtgtgtgtgagagagcgtgagtggtgtcacaggtcatcatcagaagcctga 4083

Db 1192 GTACAGCTGTGCTGGAGAGATGATGATGTGTGTCAGCGGTCAATCAGCAAGACTTGCC 1251
 Qy 4084 agtccagtccttcttctcactggtgtgacgtcacaacagatatttcaagatgattga 4143
 Db 1252 AATCTTGTCTTCTTCTTCACTGGGCTGACATCAACAGATATTTTCCGATGATTAGGC 1311
 Qy 4144 ccttcctcaatgtgaatcttctcactgaagcgaagacagatgagtcctgaattga 4203
 Db 1312 CCTCCACACTTCAAGCAATTCATGCTGACCAAGAGCAAGATGCTTACGCTTACACTT 1371
 Qy 4204 gttgagatatgcaagccctcactgagacttaaccattatgttaaacagtgattga 4263
 Db 1372 GTAGGGGGCTATGCGAGCCCTCATGAGACTTACCAATTTATGTGAAGAGTGTGGA 1431
 Qy 4264 aaggagacagcctctgaagaagcgtctgaagaaggcgatcagatcactgtcgtcaat 4323
 Db 1432 AAGGGAGCAGCAGCAAGAAATGGCTCTTAAAGAGGGGATCAGATTCCTGCAAT 1491
 Qy 4324 gggcagagctagaagagatcaacccaatgaagagctgtgtcactccttaaggaacaa 4383
 Db 1492 GGGCAAGTCTGAAAGAGATGACCCATGAAAGCTGTGCAATCTTCAAGAGGACAAAG 1551
 Qy 4384 ggcactgtcacttgatgtgtctctcttgaatgtgtcgcagaatgtaaaccaaccac 4443
 Db 1552 GGCAGCGTCAACCTCATGTGTTCTCTTGAAGTACTGCCAGAGTGAAGAGCCAGCC 1611
 Qy 4444 cctagtcacctcctcactgttgaagaatgtcactgtcctgaacttttatgctgtgt 4503
 Db 1612 ACTGGCTCCCTCTTCACTTAAACAGAGAGAGACTG-----TTGTATGCTGTGTT 1661
 Qy 4504 cagcggtgtcctcaaaacgtgaggggaatgaacacttaagttcttcttctcactg 4563
 Db 1662 GGTCCGAGAGGCTTACAGTAGGAGGCGAGA-----AAAAAAAAA 1694
 Qy 4564 aatgcttctcactgtaaacacttaacatcatttcttcttcttctgact-----ttg 4618
 Db 1695 CAGAGTGTGTTACTCACAGCAACATCATTTTCTTACTTCTGATTCATGATCATGA 1754
 Qy 4619 tgaacttaagaagaatattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4678
 Db 1755 TATACCAAAAGAAAGATATTTCATGATTAACCTGATTTATCT-CCGCAATTAAC 1813
 Qy 4679 taatgttctgtacacagggcaagaatattatcactgtaagctgtgtgtgtgtgtgt 4738
 Db 1814 TAAATTTTAAGTTCACGTCGACAA-ATTATTTATGTTCACTTGTATGAGAGAAAC 1872
 Qy 4739 agataatcctaagctaaccaa 4760
 Db 1873 AAAATGATACAAAGTTAGGCAA 1894
 RESULT 7
 AL161449/c 169660 bp DNA 09-JUL-2000
 LOCUS Human DNA sequence from clone RP11-382H24 on chromosome 9p22.1-23,
 DEFINITION complete sequence.
 ACCESSION AL161449
 VERSION AL161449.7 GI:8894259
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 169660)
 AUTHORS Sehra,H.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUL-2000) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humgeny@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Jul 1, 2000 this sequence version replaced gi:8653807.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr9>

RP11-382H24 is from the library RP11-1.2 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/>

VECTOR: pBAC3.6

This sequence is the entire insert of clone RP11-382H24. The true left end of clone RP11-664D14 is at 135950 in this sequence. The true right end of clone RP11-187K14 is at 53851 in this sequence.

Location/Qualifiers

1. 169660

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="9"

/map="P22.1-23"

/clone="RP11-382H24"

/clone_11b="RP11-1.2"

1659..1724

/note="12 copies 28 mer 92% conserved"

2068..2401

/note="LIMC4 repeat: matches 7453..7834 of consensus"

2409..2828

/note="15 copies 28 mer 62% conserved"

2429..2848

/note="7 copies 60 mer 65% conserved"

2452..2801

/note="175 copies 2 mer 61% conserved"

4621..4865

/note="MIR repeat: matches 1..261 of consensus"

5200..5294

/note="L2 repeat: matches 2578..2670 of consensus"

6031..6654

/note="LIPAL2 repeat: matches 5583..6152 of consensus"

7431..7737

/note="AluSc repeat: matches 1..307 of consensus"

7763..8103

/note="MIR1A1 repeat: matches 205..536 of consensus"

8379..8422

/note="22 copies 2 mer 76% conserved"

9774..9835

/note="LIM47 repeat: matches 6223..6284 of consensus"

10099..10533

/note="MLT1C repeat: matches 9..466 of consensus"

11109..11240

/note="AluSc repeat: matches 1..309 of consensus"

11224..11267

/note="MERSA repeat: matches 62..189 of consensus"

14993..15030

/note="MERSB repeat: matches 138..178 of consensus"

16882..16772

/note="19 copies 2 mer 97% conserved"

16759..16999

/note="MLT2D repeat: matches 1..90 of consensus"

17339..17638

/note="MIR1A1 repeat: matches 1..241 of consensus"

50925..51071

/note="AluY repeat: matches 4..304 of consensus"

repeat_region

17677..18235
/note="L2 repeat: matches 2009..2750 of consensus"

repeat_region

18882..18943
/note="MIR1J repeat: matches 411..467 of consensus"

repeat_region

18986..19291
/note="AluSc repeat: matches 1..307 of consensus"

repeat_region

19450..19635
/note="MIR repeat: matches 59..248 of consensus"

repeat_region

20230..20295
/note="53 copies 2 mer at 75% conserved"

repeat_region

21640..22131
/note="MIR1H repeat: matches 40..547 of consensus"

repeat_region

22141..22256
/note="MIR1H repeat: matches 236..352 of consensus"

repeat_region

22708..22940
/note="LIP3 repeat: matches 5401..5632 of consensus"

repeat_region

22941..23436
/note="LIP3 repeat: matches 5641..6146 of consensus"

repeat_region

26193..26290
/note="MIR repeat: matches 97..193 of consensus"

repeat_region

27837..27899
/note="L2 repeat: matches 2641..2708 of consensus"

repeat_region

28653..28739
/note="AluJo/FRAM repeat: matches 202..289 of consensus"

repeat_region

29135..29237
/note="MIR repeat: matches 97..194 of consensus"

repeat_region

29262..29788
/note="L2 repeat: matches 2198..2742 of consensus"

repeat_region

29958..30391
/note="MIR2B repeat: matches 1..444 of consensus"

repeat_region

30392..30590
/note="HERVL repeat: matches 5456..5654 of consensus"

repeat_region

30591..30631
/note="MIR2B repeat: matches 1..41 of consensus"

repeat_region

30633..31387
/note="L2 repeat: matches 1156..1991 of consensus"

repeat_region

31401..31876
/note="MIR1H repeat: matches 13..516 of consensus"

repeat_region

31936..32142
/note="L2 repeat: matches 929..1149 of consensus"

repeat_region

32790..33087
/note="AluSc repeat: matches 1..297 of consensus"

repeat_region

33089..35209
/note="LIM4 repeat: matches 4928..5049 of consensus"

repeat_region

35239..36006
/note="LIM1 repeat: matches 5352..6103 of consensus"

repeat_region

37784..37985
/note="LIMC4 repeat: matches 5695..5897 of consensus"

repeat_region

38615..38841
/note="MIR repeat: matches 3..241 of consensus"

repeat_region

38875..39054
/note="53 copies 60 mer 76% conserved"

repeat_region

38900..39039
/note="5 copies 28 mer 95% conserved"

repeat_region

39052..39352
/note="AluSc repeat: matches 1..302 of consensus"

repeat_region

39385..39681
/note="AluSc repeat: matches 1..298 of consensus"

repeat_region

40167..40470
/note="AluSc repeat: matches 5..309 of consensus"

repeat_region

40879..41165
/note="AluJo repeat: matches 1..286 of consensus"

repeat_region

44053..44310
/note="AluY repeat: matches 53..308 of consensus"

repeat_region

44543..45059
/note="MIR1H repeat: matches 64..535 of consensus"

repeat_region

45124..45229
/note="L2 repeat: matches 2160..2272 of consensus"

repeat_region

47382..47705
/note="AluJo repeat: matches 1..312 of consensus"

repeat_region

48964..49265
/note="AluSc repeat: matches 11..312 of consensus"

repeat_region

50925..51071


```

repeat_region /note="L2 repeat: matches 2579. .2750 of consensus"
51236. 51431
/note="WERB repeat: matches 9. .210 of consensus"
51507. 51613
repeat_region /note="MIR repeat: matches 53. .146 of consensus"
51614. 52055
repeat_region /note="MSTR repeat: matches 5. .426 of consensus"
52056. 52103
/note="MIR repeat: matches 3. .53 of consensus"
52372. 53211
repeat_region /note="L1MD2 repeat: matches 4972. .5856 of consensus"
53204. 53521
repeat_region /note="L1MD2 repeat: matches 4649. .4975 of consensus"
53617. 53744
/note="64 copies 2 mer aa 58% conserved"
54114. 54573
repeat_region /note="MLRIC repeat: matches 1. .466 of consensus"
55083. 55144
repeat_region /note="31 copies 2 mer ca 71% conserved"
55981. 56112
repeat_region /note="MLT11 repeat: matches 3. .136 of consensus"
57246. 57712
repeat_region /note="MLT1D repeat: matches 1. .502 of consensus"
58718. 59833
repeat_region /note="L1MB1 repeat: matches 5049. .6170 of consensus"
60653. 61120
/note="L1MA9 repeat: matches 5687. .6139 of consensus"
62690. 62872
repeat_region /note="L1MA9 repeat: matches 6130. .6308 of consensus"
64488. 64765
repeat_region /note="AlusX repeat: matches 1. .286 of consensus"
65454. 65631
repeat_region /note="L1 repeat: matches 4470. .4650 of consensus"
65633. 66246

```

```

Query Match 12.6%; Score 614.4; DB 90; Length 169660;
Best Local Similarity 99.8%; Pred. No. 1e-139;
Matches 615; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 4265 agggagcagcctctgaagcagcgtctgaagggcgatcagatcattgctgcaatg 4324
DB 146842 AGGGAGCAGCCTCTGAMAGCGACGCTGMAAGGGCGCATGAGATTCATTCGTAAG 146783
QY 4325 ggcagagctagaagagtcaccatgaagagctctgacatccttaagcgagcaaaag 4384
DB 146782 GGCAGAGCTTAAAGAGAGTCACCATGAGAGCTGTTCCTTAACGAGACAAAG 146723
QY 4385 gcactgacacttgatgctctctctgaattgctgccaagaattgaaccaaccaccc 4444
DB 146722 GCACGTGCTGCTTGAATGCTTCTCTGTAATGGCTGCCAGAAATTGAACCAACCC 146663
QY 4445 ctgactcacctcctcactgtaaaagaaagcactggtcctgacaatttttaagctgttc 4504
DB 146662 CTAGCTTCACCTCTACTGTAAGAGAAATGACATGGCTCGACAAATTTTATCTGCTTC 146603
QY 4505 agccgggtcttaaaactgtaaggggaaataaaccttaagttcttttctcatctaga 4564
DB 146602 ACCCGGGTCTTAAACTGTAGGGGGAATTAACACTTAAGTTCTTTTTCATCTAGA 146543
QY 4565 aatgcttctctactgacaacctaaccatctttcttctctctgcaatttgtaaat 4624
DB 146542 AATGCTTCTTCTACTGACACACTAACATCATTTTCTTTCTTCTTCTGCAATTTTGCAACT 146483
QY 4625 taagaagaagaatacttgtaagtgtaactcgttttatttatttgtagagatacttaatg 4684
DB 146482 TAAGAGGAAGAAATATTGTAGGTGATCTGCTTTTATTGTGAGATATCTAATCT 146423
QY 4685 tttagagtcacatgggcaagaattattacatgctaaagctggttagtataaagaagataa 4744
DB 146422 TTGTAGTACATGGCAAGAAATTTATCATCTAAGCTGCTTGTAGATTAAGAAACATTA 146363
QY 4745 ttctaaagtaacaaagaagaatgctcagtaagtaagtaagaataaataataa 4804

```

```

DB 146362 TTCTAAGCTAACCAAGAAATGCGTTCAGTAATTAAGATGAAATAATATAA 146303
QY 4805 ataagaagaatactcggggaggtttaaanaaatgctcaatttggaattaccctc 4864
DB 146302 ATAAAGAGAAATCTCGGGGAGTTTAAATAAAATGCTCATTTGGCAATCTACTCTCT 146243
QY 4865 cttcccccact 4880
DB 146242 CTCCTCCCTCCCAACT 146227

```

```

RESULT 8
AL353639/C DNA HTG 19-FEB-2001
LOCUS Homo sapiens chromosome 9 clone RP11-664D14, *** SEQUENCING IN
DEFINITION AL353639
ACCESSION AL353639
VERSION AL353639.7 GI:13016506
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

```

```

REFERENCE 1 (bases 1 to 196607)
AUTHORS Plumb, B.
TITLE Direct Submission
JOURNAL Submitted (18-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
CH10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Feb 20, 2001 this sequence version replaced gi.12964414.
COMMENT ----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: ba664D14
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 195295 bases at least Q40
Consensus quality: 195779 bases at least Q30
Consensus quality: 196025 bases at least Q20
Insert size: 196207; sum-of-ctrls
Insert size: 186500; 3.2% error; agarose-fp
Quality coverage: 7.40x in Q20 bases; sum-of-ctrls quality
coverage: 7.94x in Q20 bases; agarose-fp
-----

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

1 5394: contig of 5394 bp in length
* 1 5395 5494: gap of 100 bp
* 5495 21865: contig of 16371 bp in length
* 21866 21965: gap of 100 bp
* 21966 183547: contig of 161582 bp in length
* 183548 183647: gap of 100 bp
* 183648 194492: contig of 10845 bp in length
* 194493 194592: gap of 100 bp
* 194593 196607: contig of 2015 bp in length.

```

FEATURES

```

source
1. 196607
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-664D14"
/clone_11b="RPCI-11.3"

```


Oy 3106 cgaacaggctgaggctgacatcgttgggggttcagacagctgctggtccttatt 3165
 Db 2723 cgtttagggccttggtctcagcatttggagagaaacacaccccttgatgctatgatt 2782
 Oy 3166 atccatgaagttatgaagaaggagcaagcatgtaagaatgaagaatcctggtgctgagat 3225
 Db 2783 atccatgaagttatgaagaaggagcaagcatgtaagaatgaagaatcctggtgctgagat 2842
 Oy 3226 cagatccttagagatgtaagaatgtaagaatgtaagaatgtaagaatgtaagaatgtaagaat 3285
 Db 2843 cagatccttagagatgtaagaatgtaagaatgtaagaatgtaagaatgtaagaatgtaagaat 2902
 Oy 3286 gtccctgagacagagcagacagagatgctgctgacactccttaccagatgagagcccatatc 3345
 Db 2903 gccctgagacagacagacagagatgctgctgacactccttaccagatgagagcccatatc 2962
 Oy 3346 aagaagagagagatgctgac 3405
 Db 2963 cggcatgagagagatgctgac 3022
 Oy 3406 ctgagatgaagatgctgac 3465
 Db 3023 ctgagatgaagatgctgac 3082
 Oy 3466 aagaagagagatgctgac 3525
 Db 3083 aagaagagagatgctgac 3142
 Oy 3526 aagaagagagatgctgac 3585
 Db 3143 aagaagagagatgctgac 3202
 Oy 3586 ctgagacagatgctgac 3645
 Db 3203 cagagacatgctgac 3262
 Oy 3646 aagaagagagatgctgac 3705
 Db 3263 aagaagagagatgctgac 3322
 Oy 3706 ctgagacagatgctgac 3765
 Db 3323 ctgagacagatgctgac 3382
 Oy 3766 gaaata-----caggagataagaacagctgacacacacacacacacacacacacacacacac 3812
 Db 3383 gaaata-----caggagataagaacagctgacacacacacacacacacacacacacacacac 3442
 Oy 3813 ctgagacagatgctgac 3872
 Db 3443 ctgagacagatgctgac 3502
 Oy 3873 ctgagacagatgctgac 3932
 Db 3503 ctgagacagatgctgac 3562
 Oy 3933 ctgagacagatgctgac 3992
 Db 3563 ctgagacagatgctgac 3622
 Oy 3993 ctgagacagatgctgac 4052
 Db 3683 ctgagacagatgctgac 3682
 Oy 4053 ctgagacagatgctgac 4112
 Db 3683 ctgagacagatgctgac 3736
 Oy 4113 ctgagacagatgctgac 4172
 Db 3737 ctgagacagatgctgac 3796

Oy 4173 gcaagacagatgctgac 4232
 Db 3797 gcaagacagatgctgac 3856
 Oy 4233 ctgagacagatgctgac 4292
 Db 3857 ctgagacagatgctgac 3916
 Oy 4293 gaaagagagatgctgac 4352
 Db 3917 gaaagagagatgctgac 3976
 Oy 4353 gaaagagagatgctgac 4412
 Db 3977 gaaagagagatgctgac 4036
 Oy 4413 aatgctgacaga 4426
 Db 4037 agcctgggagcctga 4050

RESULT 10
 AF060539
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AF060539 3761 bp mRNA ROD 30-JUN-1998
 Mus musculus channel interacting PDZ domain protein mRNA, complete cds.
 AF060539
 AF060539.1 GI:3108056

house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

FEATURES

source

CDS

location

1..3761

/organism="Mus musculus"

/strain="BALB/c"

/db_xref="taxon:10090"

/tissue_type="cerebellum; brain"

250..2088

/note="CIRP"

/codon_start=1

/product="channel interacting PDZ domain protein"

/protein_id="AAC0148.1"

/translation="MVGGFPEKIRORYADLPGLHIIIELEKDNGLSLAGNKNDRS
RMSIFVGVINDEGPAAADGNMRIGDELLEINNOILYRSNOMASAIKTPATRYKLVF
IRNEDAVSOMAVAPPELSPSPVEDLGTTELVSSEPSVAKHLPPESSKPEPL
SOYVDDMVAEEOCKESPPSACQIQOITYSTOYSSSDSPSPAPLQSNADVT
GSGNFQAPLVDPAIPVDPATCTVPGDMHIIETISGRSGGLSTVGGKDTPLDAIV
IHEVEEGGAARDRLMAGDQIILEVNDLRSSEHEAIALRQTPQKRLVYRDA
OYRDEENLEFVLQDKTKGRGLGISTVGRSGGYFISDIVGGAADLDGRILRGO
ILSYNGEDMHASQETVATILKCVGLQLEIGLRAGSMAASRKTSONSGOHSNH
SSCRSEAPVITSLONLVGRSRSDPDKCTEERPRVETIIRLSALGISTSGGGS
PLDIPILFIIMAGVAAARTOKIKVDGRIVSINGOPLDGLSTIDVNLKNFGRRI
LOYVADPNIASINQLEIKMSAGSLGSPTRADRPDEDEOMORTAD"

BASE COUNT

1023 a 881 c 985 g 872 t

ORIGIN

OY 4602 ttctcttgcaatttgggaactaaaggggaaggaatatttgtagtgaaatcgcgtt 4661
 DB 279 tttcttcttgcaatttgggaactaaaggggaaggaatatttgtagtgaaatcgcgtt 220
 OY 4662 ttatttggaagatactcaatgcttctgtagtcacatgggcaagaattatcaatcgaag 4721
 DB 219 ttatttggaagatactcaatgcttctgtagtcacatgggcaagaattatcaatcgaag 160
 OY 4722 ctggttagataaagaagaataatcttaagctaaccaagaagaatggcttcagtaagt 4781
 DB 159 ctggttagataaagaagaataatcttaagctaaccaagaagaatggcttcagtaagt 100
 OY 4782 aggaatgaagaataaataaagaagaagaatctcgagggttttaaaaaatgac 4841
 DB 99 aggaatgaagaataaataaagaagaagaatctcgagggttttaaaaaatgac 40
 OY 4842 ctcaattggcaatctacctctctcccaaccacaact 4880
 DB 39 ctcaattggcaatctacctctctcccaaccacaact 1

RESULT 12
 LOCUS G27450 339 bp DNA STS 28-JUN-1996
 DEFINITION human STS SHGC-32204, sequence tagged site.
 ACCESSION G27450
 VERSION G27450.1 GI:1396169
 KEYWORDS STS; STS sequence; primer; sequence tagged site.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 339)
 AUTHORS Myers, R.M.
 JOURNAL Unpublished (1996)
 COMMENT

Contact: Richard M. Myers
 Stanford Human Genome Center (SHGC)
 Stanford University School of Medicine
 Department of Genetics, M-344, Stanford, CA 94305, USA
 Tel: 4157259687
 Fax: 4157259689
 Email: myers@shgc.stanford.edu
 Primer A: GGGAGAGAGAGTAGATTGC
 Primer B: TCTTAAGTTAAACCAAGAAATGCG
 STS size: 125
 PCR Profile:
 Initial incubation: 94 degrees C for 90 seconds
 Denaturation: 94 degrees C for 15 seconds
 Annealing: 62 degrees C for 23 seconds
 Polymerization: 72 degrees C for 30 seconds
 PCR Cycles: 30
 Thermal Cycler: Perkin Elmer 9600
 Protocol:
 Template: 25 ng
 Primer: each 1 uM
 dNTPs: each 200 uM
 Tag Polymerase: 0.05 units/uL
 Total Vol: 10 uL

Buffer:
 MgCl2: 2.5 mM
 KCl: 50 mM
 Tris-HCl: 20 mM
 pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from Z39533
 -- Washington University/Merck ESF sequence.
 Location/Qualifiers
 1. 339

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="9"
 STS
 primer_bind 11..135
 primer_bind complement(111..135)
 BASE COUNT 114 a 54 c 52 g 118 t 1 others
 ORIGIN

Query Match 6.8%; Score 333.2; DB 54; Length 339;
 Best Local Similarity 98.8%; Pred. No. 7.3e-71;
 Matches 335; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 4542 taagttcttcttcacatagaatgcttctcttactgacaacctaactatcttct 4601
 DB 339 TAAGTTCTNTTTCACATGAGAAATGCTTCTTACGACAACTAACATCATTTT 280
 OY 4602 ttctcttgcaatttgggaactaaagaagaataattgtagtgaaatcgcgtt 4661
 DB 279 tttcttcttgcaatttgggaactaaagaagaataattgtagtgaaatcgcgtt 220
 OY 4662 ttatttggaagatactcaatgcttctgtagtcacatgggcaagaattatcaatcgaag 4721
 DB 219 ttatttggaagatactcaatgcttctgtagtcacatgggcaagaattatcaatcgaag 160
 OY 4722 ctggttagataaagaagaataatcttaagctaaccaagaagaatggcttcagtaagt 4781
 DB 159 ctggttagataaagaagaataatcttaagctaaccaagaagaatggcttcagtaagt 100
 OY 4782 aggaatgaagaataaataaagaagaagaatctcgagggttttaaaaaatgac 4841
 DB 99 aggaatgaagaataaataaagaagaagaatctcgagggttttaaaaaatgac 40
 OY 4842 ctcaattggcaatctacctctctcccaaccacaact 4880
 DB 39 ctcaattggcaatctacctctctcccaaccacaact 1

RESULT 13
 LOCUS AC013514 147739 bp DNA HTG 13-JUN-2000
 DEFINITION Homo sapiens clone RP11-115J23, LOW-PASS SEQUENCE SAMPLING.
 ACCESSION AC013514
 VERSION AC013514.3 GI:9112399
 KEYWORDS HTG; HTGS_PHASE0.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 147739)
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
 Brown, A., Casale, A., Collins, S., Collins, S., Collins, S., Collins, S.,
 Cooke, P., DeBartolomeo, K., Dewar, K., Dominko, T., Doyle, M.,
 Ferrer, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,
 Galagan, J., Gargna, S., Grant, G., Hagos, B., Heath, A., Horton, L.,
 Howland, J., Johnson, R., Jones, C., Kan, L., Karas, A., Klein, J.,
 Lehoczkay, J., Liu, C., Locke, K., MacDonald, P., Marquis, N.,
 McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,
 Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
 Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Tesfaye, S., Tjirelli, A., Vassiliou, H., Vo, A., Wheeler, J., Wu, X.,
 Wyman, D., Ye, M.D., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (13-NOV-1999) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 13, 2000 this sequence version replaced gi:653323.

TITLE
 JOURNAL
 COMMENT

All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L1888

Center clone name: 115-J_23

* NOTE: This record contains 145 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

* 1
* 901 1000: contig of 900 bp in length
* 1001 1936: contig of 936 bp in length
* 1937 2036: gap of 100 bp
* 2037 2954: contig of 918 bp in length
* 2955 3054: gap of 100 bp
* 3055 3990: contig of 936 bp in length
* 3991 4090: gap of 100 bp
* 4091 4990: contig of 900 bp in length
* 4991 5090: gap of 100 bp
* 5091 6035: contig of 945 bp in length
* 6036 6135: gap of 100 bp
* 6136 7056: contig of 921 bp in length
* 7057 7156: gap of 100 bp
* 7157 8072: contig of 916 bp in length
* 8073 8172: gap of 100 bp
* 8173 9111: contig of 939 bp in length
* 9112 9211: gap of 100 bp
* 9212 10132: contig of 921 bp in length
* 10133 10232: gap of 100 bp
* 10233 11162: contig of 930 bp in length
* 11163 11262: gap of 100 bp
* 11263 12197: contig of 935 bp in length
* 12198 12297: gap of 100 bp
* 12298 13214: contig of 917 bp in length
* 13215 13314: gap of 100 bp
* 13315 14238: contig of 924 bp in length
* 14239 14338: gap of 100 bp
* 14339 15236: contig of 898 bp in length
* 15237 15336: gap of 100 bp
* 15337 16234: contig of 898 bp in length
* 16235 16334: gap of 100 bp
* 16335 17115: contig of 781 bp in length
* 17116 17215: gap of 100 bp
* 17216 18152: contig of 937 bp in length
* 18153 18252: gap of 100 bp
* 18253 19139: contig of 887 bp in length
* 19140 19239: gap of 100 bp
* 19240 20198: contig of 959 bp in length
* 20199 20298: gap of 100 bp
* 20299 21232: contig of 934 bp in length
* 21333 21332: gap of 100 bp
* 21333 22254: contig of 922 bp in length
* 22255 22354: gap of 100 bp
* 22355 23282: contig of 928 bp in length
* 23283 23382: gap of 100 bp
* 23383 24294: contig of 912 bp in length
* 24295 24394: gap of 100 bp
* 24395 25323: contig of 929 bp in length
* 25324 25423: gap of 100 bp

* 25424 26378: contig of 955 bp in length
* 26379 26478: gap of 100 bp
* 26479 27442: contig of 964 bp in length
* 27443 27542: gap of 100 bp
* 27543 28472: contig of 930 bp in length
* 28473 28572: gap of 100 bp
* 28573 29512: contig of 940 bp in length
* 29513 29612: gap of 100 bp
* 29613 30519: contig of 907 bp in length
* 30520 30619: gap of 100 bp
* 30620 31561: contig of 942 bp in length
* 31562 31661: gap of 100 bp
* 31662 32590: contig of 929 bp in length
* 32591 32690: gap of 100 bp
* 32691 33629: contig of 939 bp in length
* 33630 33729: gap of 100 bp
* 33730 34659: contig of 930 bp in length
* 34660 34759: gap of 100 bp
* 34760 35664: contig of 905 bp in length
* 35665 35764: gap of 100 bp
* 35765 36711: contig of 947 bp in length
* 36712 36811: gap of 100 bp
* 36812 37755: contig of 944 bp in length
* 37756 37855: gap of 100 bp
* 37856 38799: contig of 944 bp in length
* 38800 38899: gap of 100 bp
* 38900 39813: contig of 914 bp in length
* 39814 39913: gap of 100 bp
* 39914 40822: contig of 909 bp in length
* 40823 40922: gap of 100 bp
* 40923 41787: contig of 865 bp in length
* 41788 41887: gap of 100 bp
* 41888 42801: contig of 914 bp in length
* 42802 42901: gap of 100 bp
* 42902 43816: contig of 915 bp in length
* 43817 43916: gap of 100 bp
* 43917 44843: contig of 927 bp in length
* 44844 44943: gap of 100 bp
* 44944 45868: contig of 925 bp in length
* 45869 45968: gap of 100 bp
* 45969 46900: contig of 932 bp in length
* 46901 47000: gap of 100 bp
* 47001 47910: contig of 910 bp in length
* 47911 48010: gap of 100 bp
* 48011 48951: contig of 941 bp in length
* 48952 49051: gap of 100 bp
* 49052 50001: contig of 950 bp in length
* 50002 50101: gap of 100 bp
* 50102 51030: contig of 929 bp in length
* 51031 51130: gap of 100 bp
* 51131 52073: contig of 943 bp in length
* 52074 52173: gap of 100 bp
* 52174 53096: contig of 923 bp in length
* 53097 53196: gap of 100 bp
* 53197 54128: contig of 932 bp in length
* 54129 54228: gap of 100 bp
* 54229 55147: contig of 919 bp in length
* 55148 55247: gap of 100 bp
* 55248 56175: contig of 928 bp in length
* 56176 56275: gap of 100 bp
* 56276 57210: contig of 935 bp in length
* 57211 57310: gap of 100 bp
* 57311 58235: contig of 925 bp in length
* 58236 58335: gap of 100 bp
* 58336 59284: contig of 949 bp in length
* 59285 59384: gap of 100 bp
* 59385 60309: contig of 925 bp in length
* 60310 60409: gap of 100 bp
* 60410 61340: contig of 931 bp in length
* 61341 61440: gap of 100 bp
* 61441 62399: contig of 959 bp in length
* 62400 62499: gap of 100 bp
* 62500 63435: contig of 936 bp in length

ADINUKS BILLEN, B., LINTON, L., NUSBAUM, C. and LANDER, E.

```

NOTE: This record contains 152 individual
sequencing reads that have not been assembled into
contigs. Runs of N are used to separate the reads
and the order in which they appear is completely
arbitrary. Low-pass sequence sampling is useful for
identifying clones that may be gene-rich and allows
overlap relationships among clones to be deduced.
However, it should not be assumed that this clone
will be sequenced to completion. In the event that
the record is updated, the accession number will
be preserved.

1      898: contig of 898 bp in length
      899 998: gap of 100 bp
      999      1871: contig of 873 bp in length
      1872 1971: gap of 100 bp
      1972 2845: contig of 874 bp in length
      2846 2945: gap of 100 bp
      2946      3797: contig of 852 bp in length
      3798 3897: gap of 100 bp
      3898      4782: contig of 885 bp in length
      4783 4882: gap of 100 bp
      4883      5767: contig of 885 bp in length
      5768 5867: gap of 100 bp
      5868      6777: contig of 910 bp in length
      6778 6877: gap of 100 bp
      6878      7751: contig of 874 bp in length
      7752 7851: gap of 100 bp
      7852      8760: contig of 909 bp in length
      8761 8860: gap of 100 bp
      8861      9737: contig of 877 bp in length
      9738 9837: gap of 100 bp
      9838      10713: contig of 876 bp in length
      10714 10813: gap of 100 bp
      10814      11722: contig of 909 bp in length
      11723 11822: gap of 100 bp
      11823      12706: contig of 884 bp in length
      12707 12806: gap of 100 bp
      12807      13680: contig of 874 bp in length
      13681 13780: gap of 100 bp

```

*	5043	50512:	gap of	100 bp
*	50513	51417:	contlg of 905 bp	in length
*	51418	51517:	gap of	100 bp
*	51518	52413:	contlg of 896 bp	in length
*	52414	52513:	gap of	100 bp
*	52514	53414:	contlg of 901 bp	in length
*	53415	53514:	gap of	100 bp
*	53515	54425:	contlg of 911 bp	in length
*	54426	54525:	gap of	100 bp
*	54526	55433:	contlg of 898 bp	in length
*	55424	55523:	gap of	100 bp
*	55524	56431:	contlg of 908 bp	in length
*	56432	56531:	gap of	100 bp
*	56532	57442:	contlg of 891 bp	in length
*	57423	57522:	gap of	100 bp
*	57523	58446:	contlg of 904 bp	in length
*	58427	58526:	gap of	100 bp
*	58527	59455:	contlg of 909 bp	in length
*	59436	59535:	gap of	100 bp
*	59536	60438:	contlg of 903 bp	in length
*	60439	60538:	gap of	100 bp
*	60539	61434:	contlg of 896 bp	in length
*	61435	61534:	gap of	100 bp
*	61535	62433:	contlg of 889 bp	in length
*	62424	62533:	gap of	100 bp
*	62524	63400:	contlg of 877 bp	in length
*	63401	63500:	gap of	100 bp
*	63501	64400:	contlg of 900 bp	in length
*	64401	64500:	gap of	100 bp
*	64501	65339:	contlg of 899 bp	in length
*	65400	65499:	gap of	100 bp
*	65500	66385:	contlg of 886 bp	in length
*	66386	66485:	gap of	100 bp
*	66486	67338:	contlg of 873 bp	in length
*	67359	67458:	gap of	100 bp
*	67459	68359:	contlg of 901 bp	in length
*	68360	68459:	gap of	100 bp
*	68460	69340:	contlg of 881 bp	in length
*	69341	69440:	gap of	100 bp
*	69441	70326:	contlg of 886 bp	in length
*	70327	70426:	gap of	100 bp
*	70427	71332:	contlg of 906 bp	in length
*	71333	71432:	gap of	100 bp

Oy	1058	agagtgatattggaattctcttgatcagcagtaacttgacctgcacatctccctcggagaaac	1117
Db	60865	AGGATGTTATTGGAATTCCTTGATCCAGTACTTGATCTGCATATGTCTCTGGAGGAC	60866
Oy	1118	tataaccagaatctctctcggaaagacagatgagaaatacaccttcgglygacataagta	1177
Db	60805	TATTTACCCAGAAATCTCTCGCAAGACAGAGTGAATAATACCTTCGGTGACATAAGTA	60744
Oy	1178	tgaggcctgcttccttgcttactataaagatatacaccctgcgaatgcatggaacaac	1237
Db	60745	TGGGGCTCTGCTTCTGGCTTTACTATTAATGAATATACACCTGCAGAAATGCTATTGAACAC	60688
Oy	1238	aatacgaatgcgaaacacaaatagctgtagcctgaaatctcatcttaaccagaagtatatac	1297
Db	60685	AATATGAATGTGAAAACACAAATAGTGCTGACGATTCATTTACAAAGTAAGTTATAT	60622
Oy	1298	caagtcgagaactctctctctgctaccagatcagctcgtgaaagagctct	1347
Db	60625	CAATGTGCAAGAACTTCCTCTGTGCTACACCGAATCAAGCTGCAAGGGGAATCT	60576

RESULT 15			
AC013514/c			
LOCUS	AC013514	14773 bp	DNA
DEFINITION	Homo sapiens clone RP11-115J23, LOW-PASS SEQUENCE SAMPLING.		
			13-JUL-2000

ACCESSION AC013514
VERSION AC013514.3 GI:9112399
KEYWORDS HTG, HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 147739)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens, clone RP11-115J23
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 147739)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barna, N., Beckert, R., Boguski, L., Bouckgeater, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collins, A.,
Cooke, P., Dearlano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,
Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lehoczky, J., Liu, C., Locke, K., Macdonald, P., Marquis, N.,
McKen, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,
Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (13-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA.
On Jul 13, 2000 this sequence version replaced gi.6533323.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIDR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L1888
Center clone name: 115_J_23

* NOTE: This record contains 145 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1
* 901 1000: contig of 900 bp in length
* 1001 1936: contig of 936 bp in length
* 1937 2036: gap of 100 bp
* 2037 2954: contig of 918 bp in length
* 2955 3054: gap of 100 bp
* 3055 3990: contig of 936 bp in length
* 3991 4090: gap of 100 bp
* 4091 4990: contig of 900 bp in length
* 4991 5090: gap of 100 bp
* 5091 6035: contig of 945 bp in length
* 6036 6135: gap of 100 bp
* 6136 7056: contig of 921 bp in length
* 7057 7156: gap of 100 bp
* 7157 8072: contig of 916 bp in length
* 8073 8172: gap of 100 bp
* 8173 9111: contig of 933 bp in length
* 9112 9211: gap of 100 bp
* 9212 10132: contig of 921 bp in length
* 10133 10232: gap of 100 bp

10233 11162: contig of 930 bp in length
* 11163 11262: gap of 100 bp
* 11263 12197: contig of 935 bp in length
* 12198 12297: gap of 100 bp
* 12298 13214: contig of 917 bp in length
* 13215 13314: gap of 100 bp
* 13315 14238: contig of 924 bp in length
* 14239 14338: gap of 100 bp
* 14339 15236: contig of 898 bp in length
* 15237 15336: gap of 100 bp
* 15337 16234: contig of 898 bp in length
* 16235 16334: gap of 100 bp
* 16335 17115: contig of 781 bp in length
* 17116 17215: gap of 100 bp
* 17216 18152: contig of 937 bp in length
* 18153 18252: gap of 100 bp
* 18253 19139: contig of 887 bp in length
* 19140 19239: gap of 100 bp
* 19240 20198: contig of 959 bp in length
* 20199 20298: gap of 100 bp
* 20299 21232: contig of 934 bp in length
* 21233 21332: gap of 100 bp
* 21333 22254: contig of 922 bp in length
* 22255 22354: gap of 100 bp
* 22355 23282: contig of 928 bp in length
* 23283 23382: gap of 100 bp
* 23383 24294: contig of 912 bp in length
* 24295 24394: gap of 100 bp
* 24395 25323: contig of 929 bp in length
* 25324 25423: gap of 100 bp
* 25424 26378: contig of 955 bp in length
* 26379 26478: gap of 100 bp
* 26479 27442: contig of 964 bp in length
* 27443 27542: gap of 100 bp
* 27543 28472: contig of 930 bp in length
* 28473 28572: gap of 100 bp
* 28573 29512: contig of 940 bp in length
* 29513 29612: gap of 100 bp
* 29613 30519: contig of 907 bp in length
* 30520 30619: gap of 100 bp
* 30620 31561: contig of 942 bp in length
* 31562 31661: gap of 100 bp
* 31662 32590: contig of 929 bp in length
* 32591 32690: gap of 100 bp
* 32691 33623: contig of 939 bp in length
* 33630 33729: gap of 100 bp
* 33730 34659: contig of 930 bp in length
* 34660 34759: gap of 100 bp
* 34760 35664: contig of 905 bp in length
* 35665 35764: gap of 100 bp
* 35765 36711: contig of 947 bp in length
* 36712 36811: gap of 100 bp
* 36812 37755: contig of 944 bp in length
* 37756 37855: gap of 100 bp
* 37856 38799: contig of 944 bp in length
* 38800 38899: gap of 100 bp
* 38900 39813: contig of 914 bp in length
* 39814 39913: gap of 100 bp
* 39914 40822: contig of 909 bp in length
* 40823 40922: gap of 100 bp
* 40923 41787: contig of 865 bp in length
* 41788 41887: gap of 100 bp
* 41888 42801: contig of 914 bp in length
* 42802 42901: gap of 100 bp
* 42902 43816: contig of 915 bp in length
* 43817 43916: gap of 100 bp
* 43917 44843: contig of 927 bp in length
* 44844 44943: gap of 100 bp
* 44944 45868: contig of 925 bp in length
* 45869 45968: gap of 100 bp
* 45969 46900: contig of 932 bp in length
* 46901 47000: gap of 100 bp
* 47001 47910: contig of 910 bp in length

Fri Jul 13 15:00:15 2001

us-09-502-698-3.rge

Page 25

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2001, 18:58:57 ; Search time 284.56 Seconds

(without alignments)
10768.063 Million cell updates/sec

Title: US-09-502-698-3

Perfect score: 4880
Sequence: 1 ccggcgccggcgacagtg.....tctctcccaaccaact 4880

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq.0601:*

- 1: /SIDSI/gcgdata/geneseq/geneseq/NA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq/NA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq/NA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq/NA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq/NA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq/NA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq/NA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq/NA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq/NA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq/NA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq/NA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq/NA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq/NA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq/NA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq/NA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq/NA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq/NA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq/NA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq/NA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq/NA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4880	100.0	4880	20	AAx29908
2	4662.6	95.5	6540	20	AAx29909
3	4525.6	92.7	7431	21	AAx236453
4	4436.2	90.9	6750	20	AAx29910
5	2825	57.9	3188	21	AAx47424
6	2781.8	57.0	2819	20	AAx29960
7	2722.2	55.8	6666	20	AAx29911
8	1433.4	29.4	1681	20	AAx252869
9	854.4	17.5	2067	19	AAV69613
10	753.4	15.4	755	20	AAx29969
11	730.4	15.0	865	20	AAx29970

12	730.4	15.0	959	20	AAx29971	PDZ domain-contains PDZ domain-contains
13	563.6	11.5	1776	20	AAx29967	DNA encoding the h
14	540.2	11.1	5836	20	AAx86366	Human secreted pro
15	524	2.5	229	21	AAx27886	Human biallelic po
16	123	2.5	125	19	AAx1808	Human biallelic po
17	123	2.5	125	19	AAx12075	EST clone GP107
18	122	2.5	361	20	AAV88388	Human biallelic po
19	121.4	2.5	125	19	AAx1809	Human biallelic po
20	108.6	2.2	936	22	AAx38252	Oligonucleotide D1
21	108.6	2.2	936	22	AAx58257	Oligonucleotide D1
22	108.6	2.2	936	22	AAx58257	Oligonucleotide D1
23	108.6	2.2	936	22	AAx58259	Oligonucleotide D2
24	108.6	2.2	936	22	AAx58262	Oligonucleotide D2
25	108.6	2.2	938	22	AAx58255	Oligonucleotide D1
26	103.4	2.1	936	22	AAx58252	Oligonucleotide D1
27	103.4	2.1	936	22	AAx58254	Oligonucleotide D1
28	103.4	2.1	936	22	AAx58257	Oligonucleotide D1
29	103.4	2.1	936	22	AAx58259	Oligonucleotide D1
30	103.4	2.1	936	22	AAx58262	Oligonucleotide D2
31	103.4	2.1	938	22	AAx58255	Oligonucleotide D1
32	79.2	1.6	214	21	AAx03452	Human secreted pro
33	66	1.4	8043	16	AAx05924	Human protein tyro
34	66	1.4	8119	21	AAx61840	DNA encoding a hum
35	65.4	1.3	2963	18	AAx793775	PSD-93 coding sequ
36	63.6	1.3	1000	17	AAx18384	Human Fas-associat
37	63.6	1.3	1830	17	AAx18383	Human Fas-associat
38	63	1.3	2534	21	AAx01182	PDZ domain-compris
39	62	1.3	1060	21	AAx98033	Human colon cancer
40	62	1.3	2249	22	AAx24178	Human secreted pro
41	62	1.3	2308	21	AAx47426	Sequence encoding
42	62	1.3	2571	21	AAx65028	Membrane-bound pro
43	62	1.3	2571	22	AAx92077	Human Prol136 cDNA
44	62	1.3	2571	22	AAx44174	Human Prol136 (UNQ
45	62	1.3	5367	22	AAx24155	Human secreted pro

ALIGNMENTS

AAx29908	AAx29908 standard; DNA; 4880 BP.
AC	AAx29908;
XX	06-JUL-1999 (first entry)
XX	Clone 38-2-1 encoding protein containing PDZ domain.
DE	PDZ domain: gene expression; human umbilical vascular endothelial cell;
XX	KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding;
KW	cell; proliferation disorder; cancer; ss.
XX	
OS	Homo sapiens.
XX	
PN	W09907846-A1.
PD	18-FEB-1999.
XX	
PF	12-AUG-1998; 98WO-JP03603.
XX	
PR	19-JUN-1998; 98JP-0189944.
XX	12-AUG-1997; 97JP-0230356.
PA	(CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
PI	Funahashi S, Miyata S;
XX	
DR	WPI: 1999-167423/14.
XX	P-PSDB; AAx04730, AAx04731.
PT	Protein containing PDZ domain, whose expression is enhanced by TNF stimulation - plays an important role in protein/protein

OY	1921	gggagtgvggagtcvggtcaagcaatvggaagtvgatvgagvggagtcatttccaacaacatgctc	1960
Db	1921	ggagatvgvggagcttcvgctcaagcaaatvggaagatvgatvgagvggagcttctccaacaacatctc	1960
OY	1981	ctgagaaatgaatccagctctgccaanaatvgaaactctgnaaccctgvgagatcagaaatcgttaag	2040
Db	1981	ctgagaaatgaatccagctctgccaanaatvgaaactctgnaaacctcvgagatcagaaatcgttaag	2040
OY	2041	ggacccagtcagatcagagtcagagccagagaaagctccatctgtgcagtgtgccccaccc	2100
Db	2041	ggacccagtcagatcagagtcagagccagagaaagctcccatctgtgcagtgtgccccaccc	2100
OY	2101	ccctcttcagagctcttgccgnaaatgggttagtgtgtccacacagtcactctgcagcaanaatc	2160
Db	2101	ccctctctcagagctcttgccgnaaatgggttagtgtgtccacacagtcactctgcagcaanaatc	2160
OY	2161	tcacagaaatctvggacaagaagagatvgagctctgtgttcaacgctggaanaaatatcacagagagct	2220
Db	2161	tcacagaaatctvggacaagaagagatvgagctctgtgttcaacgctggaanaaatatcacagagagct	2220
OY	2221	tatvgaaaccttaaaagctcgagctcgcatatgatatgaaactvgagaaagatccatagtgtgttg	2280
Db	2221	fatvgaaaccttaaaagctcgagctcgcatatgatatgaaactvgagaaagatccatagtgtgttg	2280
OY	2281	ggccttaagctctctgctvggacaagaagccgatccacagatagatgtcttcatagtctggagat	2340
Db	2281	ggccttaagctctctgctvggacaagaagccgatccacagatagatgtgtcttcatagtctggagat	2340
OY	2341	gacccaaatvgagagctvgagaaagaaagatvgctgcagatctgcaaatctgcagatgagctctcaag	2400
Db	2341	gacccaaatvgagagctvgagaaagaaagatvgctgcagatctgcaaatctgcagatgagctctcaag	2400
OY	2401	atcaaatgtctcagatcttatactggaagaagaatcactcaagaatagctccatacgaataataaagt	2460
Db	2401	atcaaatgtctcagatcttatactggaagaagaatcactcaagaatagctccatacgaataataaagt	2460
OY	2461	ggccctctcaaaagtvgaaataaattttatctacgaanaataaagatvgcatcagatc	2520
Db	2461	ggccctctcaaaagtvgaaataaattttatctacgaanaataaagatvgcatcagatc	2520
OY	2521	gtatgtctctggaagaatvgagatgagaagaccttggtcccttcacactcgaanaaatctccaanaatag	2580
Db	2521	gtatgtctctggaagaatvgagatgagaagaccttggtcccttcacactcgaanaaatctccaanaatag	2580
OY	2581	gagaacagagcccaaatgtcttactactctcgtatcgtacgctgtgagactcagttcatcttcaaat	2640
Db	2581	gagaacagagcccaaatgtcttactactctcgtatcgtacgctgtgagactcagttcatcttcaaat	2640
OY	2641	gtgcacaacatctvgagagcttcccaagagatcaggggggtttvggttatgtgtcaacgccaagaa	2700
Db	2641	gtgcacaacatctvgagagcttcccaagagatcaggggggtttvggttatgtgtcaacgccaagaa	2700
OY	2701	gatacacatctvgagatgcatacaaaagagacttcaacagaaacatvggggtgtagcgccagagat	2760
Db	2701	gatacacatctvgagatgcatacaaaagagacttcaacagaaacatvggggtgtagcgccagagat	2760
OY	2761	ggagcagatccaagaagtccggagatcagatacagctgtgtagatgatagtgaataatgtgtgtgttac	2820
Db	2761	ggagcagatccaagaagtccggagatcagatacagctgtgtagatgatagtgaataatgtgtgtgttac	2820
OY	2821	ccctcttgaaaaagtattatagctcttcggaagaacagccaabgatgaacgttaaaacttaacatc	2880
Db	2821	ccctcttgaaaaagtattatagctcttcggaagaacagccaabgatgaacgttaaaacttaacatc	2880
OY	2881	catgtctgagaataccagaagcttcccgagctgttcccttcacagccagctcggtgacagctgagaa	2940
Db	2881	catgtctgagaataccagaagcttcccgagctgttcccttcacagccagctcggtgacagctgagaa	2940
OY	2941	aaaaaagaacagcttcccgagcttctgtatgcttcccaagatctgtgcttccccaagaacggaagctc	3000
Db	2941	aaaaaagaacagcttcccgagcttctgtatgcttcccaagatctgtgcttccccaagaacggaagctc	3000

QY	3001	atccgaataaagcagatcatcataacccaagcaatttttgctcttcgataccctcgacaactcgc	3060
Db	3001	atccgaaataaagcagatcatcataacccaagcaatttttgctcttcgataccctcgacaactcgc	3060
QY	3061	cccatatcccttcgctcgcgaaacacacatctcgagatttcccaaggcgcgaaacggcgctggcg	3120
Db	3061	cccatatcccttcgctcgcgaaacacacatctcgagatttcccaaggcgcgaaacggcgctggcg	3120
QY	3121	ctgagcactcgtctgggggttcaagacacgcgtcgtggtgccttatatccagaaagttaat	3180
Db	3121	ctgagcactcgtctgggggttcaagacacgcgtcgtggtgccttatatccagaaagttaat	3180
QY	3181	gaagaaagagagacagatgttaaaga ttgaaagactctggcgagatccagaaattcttagagtg	3240
Db	3181	gaagaaagagagacagatgttaaaga ttgaaagactctggcgagatccagaaattcttagagtg	3240
QY	3241	aatggaattgactctgaggaagcgccacacatgaaagcaatcaatgctcttcagacagacg	3300
Db	3241	aatggaattgactctgaggaagcgccacacatgaaagcaatcaatgctcttcagacagacg	3300
QY	3301	cccaagagagatgctgcgcgaaacactctcagaaagtctgagccccatataaaggagagagtg	3360
Db	3301	cccaagagagatgctgcgcgaaacactctcagaaagtctgagccccatataaaggagagagtg	3360
QY	3361	tgtaacacccctcaatatcttgacgtcgacgaagaagcccggaagaaagccttgagattagatt	3420
Db	3361	tgtaacacccctcaatatcttgacgtcgacgaagaagcccggaagaaagccttgagattagatt	3420
QY	3421	gttgtaaaagaaacgcgatctcgtagatatttggtcagacatctgtcaaaaggagaattgca	3480
Db	3421	gttgtaaaagaaacgcgatctcgtagatatttggtcagacatctgtcaaaaggagaattgca	3480
QY	3481	gatacccgatgaaacacgcgatcccgaggagacacagatacttgctaaagtgggaagacgtt	3540
Db	3481	gatacccgatgaaacacgcgatcccgaggagacacagatacttgctaaagtgggaagacgtt	3540
QY	3541	cgtaatgcctcccaagaagcgtgttcgcgctttgctaaagtgttccctagcgacagtaaac	3600
Db	3541	cgtaatgcctcccaagaagcgtgttcgcgctttgctaaagtgttccctagcgacagtaaac	3600
QY	3601	ttgaaagtttgaaaaatccaagactctgtccatctccaatccaagaaaggagagccatccaaac	3660
Db	3601	ttgaaagtttgaaaaatccaagactctgtccatctccaatccaagaaaggagagccatccaaac	3660
QY	3661	agccaggtgagtgaaaggcagcgtctctcttccacttttccactctcgtatccaagta	3720
Db	3661	agccaggtgagtgaaaggcagcgtctctcttccacttttccactctcgtatccaagta	3720
QY	3721	ttctagttacactggaaagttaactcaaaagaagtgcatactggcaatccaaatccaggatata	3780
Db	3721	ttctagttacactggaaagttaactcaaaagaagtgcatactggcaatccaaatccaggatata	3780
QY	3781	agaaacagtcgaaatgaaaaaaggcccttactgatactcttggaatcatagacatctgcggagga	3840
Db	3781	agaaacagtcgaaatgaaaaaaggcccttactgatactcttggaatcatagacatctgcggagga	3840
QY	3841	gttaggcagcccaacttggtgtgattgtgcctatatattatgcaayaaatgcacccaactggaagt	3900
Db	3841	gttaggcagcccaacttggtgtgattgtgcctatatattatgcaayaaatgcacccaactggaagt	3900
QY	3901	gaaagacagaccccaaaaacccagagtctggggaataagattgtccacacatctggtggcacatcc	3960
Db	3901	gaaagacagaccccaaaaacccagagtctggggaataagattgtccacacatctggtggcacatcc	3960
QY	3961	actgagggcactgactcaaccccaagcagatgtaaacctctgaaaaatgcatctgctgcattat	4020
Db	3961	actgagggcactgactcaaccccaagcagatgtaaacctctgaaaaatgcatctgctgcattat	4020
QY	4021	gaaatgagcaggtggtctgttggaaggagccgtagtggtgtgtacagagtcacatcatcaggagcct	4080
Db	4021	gaaatgagcaggtggtctgttggaaggagccgtagtggtgtgtacagagtcacatcatcaggagcct	4080
QY	4081	gaaagttccagttcttcttccactctggcgtagcgtcaaacacagatatttccagagatgattata	4140

Db	4081	gcaagctccagctctctcttcttcactctggcggaagctcaaccagatatacttcagatgattta	4140
Oy	4141	ggaccctcccaatgtaagctctattacacctaaagacgagacacagatggttaagcttcagct	4200
Db	4141	ggaccctcccaatgtaagctctattacacctaaagacgagacacagatggttaagcttcagct	4200
Oy	4201	atagttgagagataatgycagacccctcatgtagagacttaccatttaagtttaaaaaagtgctt	4260
Db	4201	atagttgagagataatgycagacccctcatgtagagacttaccatttaagtttaaaaaagtgctt	4260
Oy	4261	gcaaaagagagacagcctctctgtagagacggaacgtctcgaaaaaagggagacacagatcgtctc	4320
Db	4261	gcaaaagagagacagcctctctgtagagacggaacgtctcgaaaaaagggagacacagatcgtctc	4320
Oy	4321	aatgtagagaggtctctagaaagagatcacaccatgaaagagcgtctctgacatctcttaaaagagaca	4380
Db	4321	aatgtagagaggtctctagaaagagatcacaccatgaaagagcgtctctgacatctcttaaaagagaca	4380
Oy	4381	aaagagacgtctacacttgaatgctctctctctctgaattgctgcgcagaaatgtaaaccaacca	4440
Db	4381	aaagagacgtctacacttgaatgctctctctctctgaattgctgcgcagaaatgtaaaccaacca	4440
Oy	4441	accctctgcacactcctctatctgtaaaagagaagcactggctcctgcacaattttatgctgt	4500
Db	4441	accctctgcacactcctctatctgtaaaagagaagcactggctcctgcacaattttatgctgt	4500
Oy	4501	gtctacagcgggtctctcaaaaactgtaaggggggaaataacacttaagcttctctttctcaac	4560
Db	4501	gtctacagcgggtctctcaaaaactgtaaggggggaaataacacttaagcttctctttctcaac	4560
Oy	4561	tagaaatgctcttcccttactctgacaacactaaacatcatcttctcttctctctgcatttg	4620
Db	4561	tagaaatgctcttcccttactctgacaacactaaacatcatcttctcttctctctgcatttg	4620
Oy	4621	aacttaagaaggagaaatatttctgtaagttgaattccgttttattattcttgtagagatatcta	4680
Db	4621	aacttaagaaggagaaatatttctgtaagttgaattccgttttattattcttgtagagatatcta	4680
Oy	4681	atgtttctgtctgcacatgycgacaagaaattatatacgtctgaagctgtgttagataaaagaag	4740
Db	4681	atgtttctgtctgcacatgycgacaagaaattatatacgtctgaagctgtgttagataaaagaag	4740
Oy	4741	ataactctaaagcttaaccaaagaataatgctcttcgaatgaatttagataaanaatganaata	4800
Db	4741	ataactctaaagcttaaccaaagaataatgctcttcgaatgaatttagataaanaatganaata	4800
Oy	4801	taaaataaagaagaanaaatctcggggggtttaaaaaaatgctcaatttggaatctaac	4860
Db	4801	taaaataaagaagaanaaatctcggggggtttaaaaaaatgctcaatttggaatctaac	4860
Oy	4861	tcctctcccaaccccaaacat 4880	
Db	4861	tcctctcccaaccccaaacat 4880	

RESULT	2
AAAX29909	
ID	AAAX29909 standard; DNA; 6540 BP.
XX	
AC	AAAX29909;
XX	
DT	06-JUL-1999 (first entry)
XX	
DE	Clone 38-2-1a encoding protein containing PDZ domain.
XX	
KW	PDZ domain; gene expression; human umbilical vascular endothelial cell;
KW	HVVEG; stimulation; tumour necrosis factor; TNF; protein binding;
KM	cell; proliferation disorder; cancer; ss.
XX	
OS	Homo sapiens.
XX	
XX	
PN	WO9907846-A1.

XX 18-FEB-1999.
PD
XX
XX PF 12-AUG-1998; 98MO-JP03603.
XX
XX PR 19-JUN-1998; 98JP-0189944.
PR 12-AUG-1997; 97JP-0230356.
XX
XX
PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
PI Funahashi S, Miyata S;
XX
XX WPI: 1999-167423/14.
DR P-PSDB: AAY04732.
XX

Protein-containing PDZ domain, whose expression is enhanced by TNF stimulation - plays an important role in protein/protein interactions and is used for screening for proteins for use in treatment of cell proliferation disorders such as cancer

PS Example 7; Page 183-201; 240pp; Japanese.
 vv

CC This sequence represents clone 38-2-1a which encodes a new protein
CC containing PDZ domains whose expression in human umbilical vascular
CC endothelial cells (HUEV) is enhanced by stimulation with tumour
CC necrosis factor (TNF) alpha. The new protein is used to identify
CC proteins which bind to it (particularly to the PDZ domains) and then
CC genes encoding them, for use in the treatment of cell proliferation
CC disorders such as cancer.

Sequence 6540 BP; 1990 A; 1330 C; 1590 G; 1630 T; 0 other;

Query Match	95.5%;	Score 4662.6;	DB 20;	Length 6540;
Query Match	95.5%;	Score 4662.6;	DB 20;	Length 6540;

Matches 4665; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY	212	aggtaaatctgcatatcaacttcttcttgaggaaatcacaaagatcttgctggaataactttaaag	271
Db	1872	aagcaaaagagcatcaactcttctcgggaaataacacaaagatctgggtgaatactttaaag	1931
OY	272	aactgcctataagatgacaaatgctgctgtctgtcgaacatgctgcacccacacccacat	331
Db	1932	aactgcctataagatgacaaatgctgctgtctgtcgaacatgctgcacccacacccacat	1991
OY	332	cagaaattggagaaagctctgaccttctatgcatattgagcttaacaaagaaagctccacgtacac	391
Db	1992	cagaaattggagaaagctctgaccttctatgcatattgagcttaacaaagaaagctccacgtacac	2051
OY	392	taagtgagcttcatactcgtgtctacacgaacagagatccaaatgctgtgcgaatgacatgctgg	451
Db	2052	taagtgagcttcatactcgtgtctacacgaacagagatccaaatgctgtgcgaatgacatgctgg	2111
OY	452	gtccagaaatacagaagaagaggtttccaagacacttttgcaatctgaggaaagctgcatctaacaca	511
Db	2112	gtccagaaatacagaagaagaggtttccaagacacttttgcaatctgaggaaagctgcatctaacaca	2171
OY	512	taagagctcgagaaagagggagcaaaagagacttggttttaacatcttatacatatacagaatccaa	571
Db	2172	taagagctcgagaaagagggagcaaaagagacttggttttaacatcttatacatatacagaatccaa	2231
OY	572	ttgtatccagcaagcacctgtgatatataatctcgttcttcttgctgtgcctgaggaacttctgaa	631
Db	2232	ttgtatccagcaagcacctgtgatatataatctcgttcttcttgctgtgcctgaggaacttctgaa	2291
OY	632	aggaatgacagactctctcctcgtgtcacagcaactcatgcttgtaaacagatgttcaactctgaa	691
Db	2292	aggaatgacagactctctcctcgtgtcacagcaactcatgcttgtaaacagatgttcaactctgaa	2351
OY	692	acagacagctcttgaggaagctcgtataagaagcaactgaaggagacacccgtccaggaactctgaa	751
Db	2352	acagacagctcttgaggaagctcgtataagaagcaactgaaggagacacccgtccaggaactctgaa	2411

OY	752	taagaaftvgctaaagccttaaccccttcacacagaaagagltatgtctctgctaaagag	811
Db	2412	taagaagtgtcgaagcctttacccttccttcacacagaaagagttatgtctctgctaaagag	2471
OY	812	attccttctctaacccacacacactcctctgtgaggaagcaggtgtgctgtgacaaacccct	871
Db	2472	attccttctcttaaccacacacactcctctgtgaggaagcaggtgtgctgtgacaaacccct	2531
OY	872	tcagagctgacttgctctgtgtgagcaaaatgatgtctgaacttaagatgaatccacat	931
Db	2532	tcagagctgacttgctctgtgtgagcaaaatgatgtctgaacttaagatgaatccacat	2591
OY	932	ttgagctcccaactcctctcgaaaatgacagacatcactcactcaacgcctcatttat	991
Db	2592	ttgagctcccaactcctctcgaaaatgacagacatcactcactcactcaacgcctcatttat	2651
OY	992	ctcttcactggaagtctctgtgtgagtgaggtcctgtgaacatgttctctcccttcacatctc	1051
Db	2652	ctcttcactggaagtcctctgtgtgagtgaggtcctgtgaacatgttctctcccttcacatctc	2711
OY	1052	cctccaaagaaatgtatctgaaaataatctctgtgatacagatcactgtatctgacatgtctcgtg	1111
Db	2712	cctccaaagaaatgtatctgaaaataatctctgtgatacagatcactgtatctgacatgtctcgtg	2771
OY	1112	aggaacataatacccaagaaatctcctcgtgaaaagacagaaatgagaatacacctcgtgtgaca	1171
Db	2772	aggaacataatacccaagaaatctcctcgtgaaaagacagaaatgagaatacacctcgtgtgaca	2831
OY	1172	taagaataggggcctgctctcgtgcttactatataatgatacacaacaccttgaaatgtatgt	1231
Db	2832	taagaataggggcctgctctcgtgcttactatataatgatacacaacaccttgaaatgtatgt	2891
OY	1232	aaacaaataatgaaatgagaaaacaacaataagatgagaaatgaaatcattatccaagaatgaa	1291
Db	2892	aaacaaataatgaaatgagaaaacaacaataagatgagaaatgaaatcattatccaagaatgaa	2951
OY	1292	ttatacaaatgtcagaagaatctctctctgtgtgtctacacgaattcacgtggaagaagctctgagc	1351
Db	2952	ttatacaaatgtcagaagaatctctctctgtgtgtgtctacacgaattcacgtggaagaagctctgagc	3011
OY	1352	aacgtcgttgaaacagaagctctccctgtgacgtcgtgaatgtctgaagtgtcattatgctcaaaatgtat	1411
Db	3012	aacgtcgttgaaacagaagctctccctgtgacgtcgtgaatgtctgaagtgtcattatgctcaaaatgtat	3071
OY	1412	ctaaagaagaatctcttgaaaaggaactatataatgacaagaagcgaattctgaactcagaatgaa	1471
Db	3072	ctaaagaagaatctcttgaaaaggaactatataatgacaagaagcgaattctgaactcagaatgaa	3131
OY	1472	cagttagatgtcataataagaatgtcgtctgtggaatgtacgtctcgaaagcattatccaatgaaatgtg	1531
Db	3132	cagttagatgtcataataagaatgtcgtcgtggaatgtacgtctcgaaagcattatccaatgaaatgtg	3191
OY	1532	ccattagtcgagaaatgagcccggaattgtccaattgaggaatcgtcatctgtccattatgaagaagt	1591
Db	3192	ccattagtcgagaaatgagcccggaattgtccaattgaggaatcgtcatctgtccattatgaagaagt	3251
OY	1592	ctacacataagatgaaaccaaatgcccaggaacagagctatgtgtgagaagaacattctcccatgt	1651
Db	3252	ctacacataagatgaaaccaaatgcccaggaacagagctatgtgtgagaagaacattctcccatgt	3311
OY	1652	ggccctgacataaaaaataactatctatgtcgtcctgcagaaaacatttggaagaagtccaataaagct	1711
Db	3312	ggccctgacataaaaaataactatctatgtcgtcctgcagaaaacatttggaagaagtccaataaagct	3371
OY	1712	tgggaacaacaatctggaagaagtaatgtgcacatgtgatactttctctcaatacacatgtgcagag	1771
Db	3372	tgggaacaacaatctggaagaagtaatgtgcacatgtgatactttctctcaatacacatgtgcagag	3431
OY	1772	acatttccgaatattccagagccggaaggaaggaaggaatgaaagaagaagcgaactccaanaa	1831
Db	3432	acatttccgaatattccagagccggaaggaaggaaggaatgaaagaagaagcgaactccaanaa	3491
OY	1832	cagcatalatgaatctggaataacagcccacagcgggtgtgaaactctcgagaggaacaacaagaacat	1891

Db	3492	cagacatataagcaattggaataacagccacgagcggaatggaacattcgagagagacaacaaacaaat	3551
Qy	1892	ccttaagagcatcaagatctgttcgttggaagagagagatggggagcttcggctcaaaacaaatggaag	1951
Db	3552	ccttaagagcatcaagatctgttcgttggaagagagagatggggagcttcggctcaaaacaaatggaag	3611
Qy	1952	tgaataggggagcatcttcatacnaaaatgctcttgaaagataatgctccagcttgccaaaaatgga	2011
Db	3612	tgaataggggagcatcttcatacnaaaatgctcttgaaagataatgctccagcttgccaaaaatgga	3672
Qy	2012	ccttgaaaaacctggaagatagaatcgttagaggacccagtcagtcagagatcagaagccaga	2072
Db	3672	ccttgaaaaacctggaagatagaatcgttagaggacccagtcagtcagagatcagaagccaga	3732
Qy	2072	aggagcccatcttgaaagtggtcccccaacccctcttcagaccccttgccagaaatgggttaag	2132
Db	3732	aggagcccatcttgaaagtggtcccccaacccctcttcagaccccttgccagaaatgggttaag	3792
Qy	2132	atcacacacagtcatactgccaagacnaaaatctcacaaagatgtagaacaagagagatgaattg	2192
Db	3792	atcacacacagtcatactgccaagacnaaaatctcacaaagatgtagaacaagagagatgaattg	3852
Qy	2192	gttaacagctggaaaaataatacagaagagcggttaatggaaccttaaaagcgagctgcataatga	2252
Db	3852	gttaacagctggaaaaataatacagaagagcggttaatggaaccttaaaagcgagctgcataatga	3912
Qy	2252	ttagaacctggagaagagatcagaatggttttggtgacctaaagtctgtctgggaacaaagccggt	2312
Db	3912	ttagaacctggagaagagatcagaatggttttggtgacctaaagtctgtctgggaacaaagccggt	3972
Qy	2312	ccaagatagaatgtcttcataagtaggggaatgataccaaatggagctgccaagaaaaagatgtgct	2372
Db	3972	ccaagatagaatgtcttcataagtaggggaatgataccaaatggagctgccaagaaaaagatgtgct	4032
Qy	2372	gatttcgaaattgccaagatggaagctcttcagaagatcaatggtcagatatttaattggaagaagctc	2432
Db	4032	gatttcgaaattgccaagatggaagctcttcagaagatcaatggtcagatatttaattggaagaagctc	4092
Qy	2432	atcagaatagcttcatacatacttaaaatggtcccttccttaagaggaataaatctttatca	2492
Db	4092	atcagaatagcttcatacatacttaaaatggtcccttccttaagaggaataaatctttatca	4152
Qy	2492	gaaataaagaatgccaagtgaatcagaatggcgatagctcctgggaatggcagatagaacctttgc	2552
Db	4152	gaaataaagaatgccaagtgaatcagaatggcgatagctcctgggaatggcagatagaacctttgc	4212
Qy	2552	ctcttaactcagaaaatctccaataaataagagagacagagccaactgttaactactctgag	2612
Db	4212	ctcttaactcagaaaatctccaataaataagagagacagagccaactgttaactactctgag	4272
Qy	2612	cagcgctggagacctgaattcatttaaaaaatgtgcaaaactcggggagcttcccaagatcacag	2672
Db	4272	cagcgctggagacctgaattcatttaaaaaatgtgcaaaactcggggagcttcccaagatcacag	4332
Qy	2672	ggggcttggtgcatctgcatcacgcaaaagaatatacaactcagtgagatcacaataaagagct	2732
Db	4332	ggggcttggtgcatctgcatcacgcaaaagaatatacaactcagtgagatcacaataaagagct	4392
Qy	2732	taaacagagcatggggtagcagaccacgggaatggaagcatcctaaagtccgggaatcgaataatcgg	2792
Db	4392	taaacagagcatggggtagcagaccacgggaatggaagcatcctaaagtccgggaatcgaataatcgg	4452
Qy	2792	ctgtgaatagatgaaatctgtgtgtgtgttaacctattgaaagaatttatagctcttcgaaga	2852
Db	4452	ctgtgaatagatgaaatctgtgtgtgtgtgttaacctattgaaagaatttatagctcttcgaaga	4512
Qy	2852	cagcaaaagaatgacggttaaacttaacatcacaatgcttagaataccaagatctcccaagcgctgtc	2912
Db	4512	cagcaaaagaatgacggttaaacttaacatcacaatgcttagaataccaagatctcccaagcgctgtc	4572
Qy	2912	cttcagcgagctgggtgcagcgagtggaagaaaaaagacagcttcccggtcttcgatgtgctc	2972

Db	4572	ctcaagcagctcgtgycagcagcgtgycgagaaanaagaaacagctcccaagctctcgtatgctc	4631
Oy	2972	cacacgctctgctctccccaagacccgagatccaatccgaaatacaagcagatcatcaacaacag	3031
Db	4632	cacacgctctgctctccccaagacccgagatccaatccgaaatacaagcagatcatcaacaacag	4692
Oy	3032	caattctgctctcgatccacgcgaacctgcgcccattaccctgcgtcgcgaaanaacatcgt	3091
Db	4692	caattctgctctcgatccacgcgaacctgcgcccattaccctgcgtcgcgaaanaacatcgt	4751
Oy	3092	agattccaaagggcgaaacagggctcgccctcgagatctggygggttcaagacacgtgc	3151
Db	4752	agattccaaagggcgaaacagggctcgccctcgagatctggygggttcaagacacgtgc	4811
Oy	3152	tgggtgctcttatcatcatcgaagaagtttatbgaagaagagcagcatgataaagtctgaagac	3211
Db	4812	tgggtgctcttatcatcatcgaagaagtttatbgaagaagagcagcatgataaagtctgaagac	4871
Oy	3212	tcgycgctbgaagatcagaatcttaagagtgaaatgaaatgtaacttggaggaagccacatag	3271
Db	4872	tcgycgctbgaagatcagaatcttaagagtgaaatgaaatgtaacttggaggaagccacatag	4931
Oy	3272	atggaagacatcaaatgctctbgaagacgaagcccaagagaaatgagccttgcaacttaacag	3331
Db	4932	atggaagacatcaaatgctctbgaagacgaagcccaagagaaatgagccttgcaacttaacag	4992
Oy	3332	atggaagccccaataaagaagagagagctgtgtgacacccctcaatactgaagctgcgaaga	3392
Db	4992	atggaagccccaataaagaagagagagctgtgtgacacccctcaatactgaagctgcgaaga	5051
Oy	3392	agccgaggaagaagagcctgaagattgaatctgtgtgtgaataaagaacgaatactgagatattg	3451
Db	5052	agccgaggaagaagagcctgaagattgaatctgtgtgtgaataaagaacgaatactgagatattg	5111
Oy	3452	tgtcagaagcatgttcaaaagggagagatctgcagatcccgatbgaagaagctgataccagggagac	3511
Db	5112	tgtcagaagcatgttcaaaagggagagatctgcagatcccgatbgaagaagctgataccagggagac	5171
Oy	3512	agataataattggtgaaatgagggaagacgcttcogtaatgctctccaaagagcgggttcgcgctt	3571
Db	5172	agataataattggtgaaatgagggaagacgcttcogtaatgctctccaaagagcgggttcgcgctt	5231
Oy	3572	tgcgaagaagtgctccctctgagacagtaaacctctggaagcttggaagaatccaagcttgtccat	3631
Db	5232	tgcgaagaagtgctccctctgagacagtaaacctctggaagcttggaagaatccaagcttgtccat	5291
Oy	3632	tcaatcttccaactctctggtatccaagatacatctgagatccaatctggaataagctgcaccaagaaga	3691
Db	5292	tcaatcttccaactctctggtatccaagatacatctgagatccaatctggaataagctgcaccaagaaga	5351
Oy	3692	tcaatcttccaactctctggtatccaagatacatctgagatccaatctggaataagctgcaccaagaaga	3751
Db	5352	tcaatcttccaactctctggtatccaagatacatctgagatccaatctggaataagctgcaccaagaaga	5411
Oy	3752	atgcatctgcaatctbgaataatacagggatctaagaacagctgnaaatgnaaaagggccctactg	3811
Db	5412	atgcatctgcaatctbgaataatacagggatctaagaacagctgnaaatgnaaaagggccctactg	5471
Oy	3812	actcaacggggaatcaagatctcgtgaggggtgagcgagcccaacttggtgatatgtcctatat	3871
Db	5472	actcaacggggaatcaagatctcgtgaggggtgagcgagcccaacttggtgatatgtcctatat	5531
Oy	3872	ttaattcgaatgatacacaacacacgtgggtctgcaagcaagaacccaataacttcgaagtgtggg	3931
Db	5532	ttaattcgaatgatacacaacacacgtgggtctgcaagcaagaacccaataacttcgaagtgtggg	5591
Oy	3932	ataggaattgtcaacatctgtgycacataccaactgagggcctgactacacccaagcagttta	3991
Db	5592	ataggaattgtcaacatctgtgycacataccaactgagggcctgactacacccaagcagttta	5651
Oy	3992	accacacggaaaaaatgatactcgtgctccaattcgaaatcgacagtggtgtcgtgagagagacgtga	4051
Db	5652	accacacggaaaaaatgatactcgtgctccaattcgaaatcgacagtggtgtcgtgagagagacgtga	5711

OY	4052	gtcggtcacaaggtgcacatcaccaagaaocgcgcgaattccagttccagttcttcttcaactgggcga	4111
Db	5712	gtcggtcacaaggtgcacacacaaacgaagagcgcgcgaattccagttcttcttcaactgggcga	5771
OY	4112	cgctacaaccagatattcttcagagagatttcgaagacctctcaatcgaagtctcttacaactag	4171
Db	5772	cgctacaaccagatattcttcagagatttcgaagacctctcaatcgaagtctcttacaactag	5831
OY	4172	agcgaaagaccagatgcttgcttgcttaagatagttgagagataatgycagccctcatgag	4231
Db	5832	agcgaaagaccagatgcttgcttgcttaagatagttgagagataatgycagccctcatgag	5891
OY	4232	acttacccattatgctttaaaacagctgttgcgaaggaagcagccctctaaagcgaagc	4291
Db	5892	acttacccattatgctttaaaacagctgttgcgaaggaagcagccctctaaagcgaagc	5951
OY	4292	tcgaaaaggggcgcatacagatcatctgctgcataatggcagagctctaaagagatcaaccatg	4351
Db	5952	tcgaaaaggggcgcatacagatcatctgctgcataatggcagagctctaaagagatcaaccatg	6011
OY	4352	aagaagcgtgtgcatactcttaaaacggaaacaaagcacttcaacttggttctctct	4411
Db	6012	aagaagcgtgtgcatactcttaaaacggaaacaaagcacttcaacttggttctctct	6071
OY	4412	gaattgctgcgcagaatctgaaccaacccaacccctatgcactctctactctaaagaga	4471
Db	6072	gaattgctgcgcagaatctgaaccaacccaacccctatgcactctctactctaaagaga	6131
OY	4472	tgcactggtccctgcacaatttttatgctgtgttcacgcgggtctccaacaactgtagggg	4531
Db	6132	tgcactggtccctgcacaatttttatgctgtgttcacgcgggtctccaacaactgtagggg	6191
OY	4532	aaataaacacttaagttctcttcttcatactcagaaagctcttcttaacgcgaacactaca	4591
Db	6192	aaataaacacttaagttctcttcttcatactcagaaagctcttcttaacgcgaacactaca	6251
OY	4592	tcaatttctcttctctctcttgcatttctgtgaacttaagagaaagaaattctgttagtg	4651
Db	6252	tcaatttctcttctctctcttgcatttctgtgaacttaagagaaagaaattctgttagtg	6311
OY	4652	aatctcgtttttatttctgtggaagatactcaatgcttctgtgatcacaatggcagaagaattat	4711
Db	6312	aatctcgtttttatttctgtggaagatactcaatgcttctgtgatcacaatggcagaagaattat	6371
OY	4712	acatgctcaagctgttgtagtaaaagaaagataattcctaagctaacaaagaaatgct	4771
Db	6372	acatgctcaagctgttgtagtaaaagaaagataattcctaagctaacaaagaaatgct	6431
OY	4772	tcagtaagtttagatgtgaaaaaatgaaaaataaaataaagaagaagaaactctgggagattta	4831
Db	6432	tcagtaagtttagatgtgaaaaaatgaaaaataaaataaagaagaagaaactctgggagattta	6491
OY	4832	aaaaaaatgctcaatttggcacaactctctctcccaaccccaact	4890
Db	6492	aaaaaaatgctcaatttggcacaactctctctcccaaccccaact	6540

RESULT	3
AAZ36453	
ID	AAZ36453 standard; DNA; 7431 BP.
XX	
AC	AAZ36453;
XX	
DT	22-FEB-2000 (first entry)
XX	
DE	DNA encoding the MMSC2 protein.
XX	
KW	Human; MMSC2; MNAC1; PDZ domain; tumour suppressor; tyrosine phosphatase;
XX	scaffolding protein; cancer; ss.
OS	Homo sapiens.
XX	

FH	Key	Location/Qualifiers
FT	CDS	57..6170
FT		/tag= a
FT		/product= "MMSC2 protein"
XX		
PN		M0998548-A1.
PD		18-NOV-1999.
XX		
PF		07-MAY-1999; 99WO-US09969.
XX		
PR		08-MAY-1998; 98US-0084740.
XX		
PA		(MYRI-) MYRIAD GENETICS INC.
XX		
P1		Bartel PL, Tavtigian SV;
XX		
DR		WPI: 2000-053077/04.
DR		P-PSDB: MAY53753.
XX		
PT		Nucleic acids and polypeptides representing human MMSC2, useful for detecting, diagnosing a predisposition to, and treating cancer -
XX		
PS		Claim 2; Page 84-93; 112pp; English.
XX		
CC		The present sequence encodes the human MMSC2 protein. The MMAC1 protein binds to MMSC2. The MMSC2 protein has 11 post-synaptic density protein, disc-large, zo-1 (PDZ) domains and one or more of these domains interacts specifically with the carboxyl terminal amino acids of MMAC1 (see AAY53754). Specifically, it appears that domain 7, 10 and 13 interact with MMAC1. Since MMSC2 contains 11 PDZ domains and interacts with MMAC1, a known tumour suppressor having a region of homology with protein tyrosine phosphatases, MMSC2 acts as a scaffolding protein in a common biological pathway with MMAC1. It is believed that the interaction between MMAC1 and MMSC2 is required for the tumour suppressor activity of MMAC1. The MMSC2 polypeptides, polynucleotides, fragments and specific or complex specific antibodies may be used for detecting cancer or a predisposition to cancer and screening for agents that may be used to treat MMSC2 and/or MMAC1 related cancer. The polypeptides and polynucleotides may also be used to treat cancer.
CC		
CC		
CC		
CC		
SO		Sequence 7431 BP: 2321 A; 1447 C; 1741 G; 1922 T; 0 other:
Query Match	92.7%; Score 4525.6; DB 21; Length 7431;	
Best Local Similarity	97.4%; Pred. No. 0;	
Matches 4655; Conservative	0; Mismatches 14; Indels 111; Gaps 1	
OY	212 aggtaatgagcataacttacttgagggaataacccaagaatgtygtgaatatcttaaaag	271
Db	1858 agttaattgcgctaactttacttcggggaataaccacaagaatgtygtgaatatcttaaaag	1917
OY	272 aactcccatagaaatgcaaatgagtgtcgtcgctcgaaactgtgccaccaccaccacct	331
Db	1918 aactgccataagaaatgcaaatgagtgtcgtcgctcgaaactgtgccaccaccaccacct	1977
OY	332 cagaatctgatgcctcggaactatgtgatatctgaactaacagaanaagctcacgtatgac	391
Db	1978 cagaatctgatgcctcggaactatgtgatatctgaactaacagaanaagctcacgtatgac	2037
OY	332 tagtgtagttcaatcggtgtcatcagaagacagagatccagttgctgvgatgactgatybcgg	451
Db	2038 tagtgtagttcaatcggtgtcatcagaagacagagatccagttgctgvgatgactgatybcgg	2097
OY	452 gtccagaagtagaaaaggtttcaagaacctttggcaatgryggaaagctgtgcatcagcaca	511
Db	2098 gtccagaagtagaaaaggtttcaagaacctttggcaatgryggaaagctgtgcatcagcaca	2157
OY	512 tagagcttgagaagaagggagcaaaagactgtgtttagacattttagattatcagatccaa	571
Db	2158 tagagcttgagaagaagggagcaaaagactgtgttttagacattttagattatcagatccaa	2217
OY	572 ttgatccagcaagcaatgattatcaattcgcttcttggtgcttcgycggcatctgtgtaa	631

Db	2218	ttgattccagacagacgtgvtatataatctgltctcttctgtgcctcgtgcgagcatctgtcggaaa	2277
Qy	6312	aggaatgacagacattctctccctcgtgtgacagacatactgttttaaacagatgttaactctggaaa	691
Db	2278	aggaatgacagacattctctccctcgtgtgacagacatactgttttttaaacagatgttaactctggaaa	2337
Qy	692	acagacagctcttgaaagacgtctagaagacatctgaagagacacggtcgaagagctctggaaa	751
Db	2338	acagacagctcttgaaagacgtctagaagacatctgaagagagacacggtcgaagagactctggaaa	2397
Qy	752	taggaagcttgtaagaccttaccctcttcaaccgaagaaggttaatgtttctgtgtaagagg	811
Db	2398	taggaagcttgtaagaccttaccctcttcaaccgaagaagagttatgtttctgtgtaagagg	2457
Qy	812	atctctcttctcaaccacacacactctctgtgtagaagacaagctctgtgcgaacacccct	871
Db	2458	atctctcttctcaaccacacacactctctgtgtagaagacaagctctgtgcgaacacccct	2517
Qy	872	tcaaggtcgtgacctgtgctctgtgtgtgagacaataatgtctgaccttaagtaagatgaaatc	931
Db	2518	tcaaggtcgtgacctgtgctctgtgtgtgagacaataatgtctgaccttaagtaagatgaaatc	2577
Qy	932	ttgagcttccatactctctccgaaaatgtgacagacatactactctcaagcctcttatct	991
Db	2578	ttgagcttccatactctctccgaaaatgtgacagacatactactctcaagcctcttatct	2637
Qy	992	ctcttcaatgcagttctctgtgtgtgtgagagcctctgaacataatgttcttcccttccataatc	1051
Db	2638	ctcttcaatgcagttctctgtgtgtgtgagagcctctgaacataatgttcttcccttccataatc	2697
Qy	1052	ctctcaagagatgttatgtgaaaatctctgtgtatccagtaactgtatctgtcaatactgtctg	1111
Db	2698	ctctcaagagatgttatgtgaaaatctctgtgtatccagtaactgtatctgtcaatactgtctg	2757
Qy	1112	aggaacatataaccagaaatctctctgtgaaaagacagatagaataacacctctgtgtgaca	1171
Db	2758	aggaacatataaccagaaatctctctgtgaaaagacagatagaataacacctctgtgtgaca	2817
Qy	1172	taagtaatgtggggccgtgtctctgtgttactataataatgacacacacacctgtgaaaatgttat	1231
Db	2818	taagtaatgtggggccgtgtctctgtgttactataataatgacacacacacctgtgaaaatgttat	2877
Qy	1232	aacacaacataatgaatgtgaaaaacacaataatgtgtgtgacatgaatctcatatcaccaagtga	1291
Db	2878	aacacaacataatgaatgtgaaaaacacaataatgtgtgtgacatgaatctcatatcaccaagtga	2937
Qy	1292	ttatatacaagtgcagaacctctctctgtgtgtctacccgaattcagctgtgaaaaggctctgagc	1351
Db	2938	ttatatacaagtgcagaacctctctctgtgtgtctacccgaattcagctgtgaaaaggctctgagc	2997
Qy	1352	accgtgtctgaaacagatgtccctctgtgtaatgtctgaatgttcaatgtcttaaaatgtat	1411
Db	2998	accgtgtctgaaacagatgtccctctgtgtaatgtctgaatgttcaatgtcttaaaatgtat	3057
Qy	1412	ctaaagaacattcttgaaaagagacattataataagacaagaagacattctagctctagaataatga	1471
Db	3058	ctaaagaacattcttgaaaagagacattataataatagacaagaagacattctagctctagaataatga	3117
Qy	1472	cagttagtgctaaataaagaatgtgctctgtggagatgactgtctcgaagaatctatctgaagagt	1531
Db	3118	cagttagtgctaaataaagaatgtgctctgtggagatgactgtctcgaagaatctatctgaagagt	3177
Qy	1532	ccatagtgagagatggcccggaattgccaattggggagacacatctgttccatcttaatggaaagt	1591
Db	3178	ccatagtgagagatggcccggaattgccaattggggagacacatctgttccatcttaatggaaagt	3237
Qy	1592	ctaacacacagtgtaacaacaatgcccagagacagactaigtcttgaaagacattctccatctg	1651
Db	3238	ctaacacacagtgtaacaacaatgcccagagacagactaigtcttgaaagacattctccatctg	3297
Qy	1652	ggcctgacataaataatctatgtgtcctgtgcagacaatttgaaagattcaaaataagct	1711

Db 3298 gccctgacataaaatlctatgtgcttcgagacatttgaagatgtcaaaataagct 3357
OY 1712 tgggacaacaacttcggaagagtaatgtgcactggaattttcttcacacatgacag 1771
Db 3358 tgggacaacaacttcggaagagtaatgtgcactggaattttcttcacacatgacag 3417
OY 1772 acaattccagaattccagagcgagagagagagagagagagagagagagagagagag 1831
Db 3418 acattccagaattccagagcgagagagagagagagagagagagagagagagagagag 3477
OY 1832 cagacataagaatttggaaataagccagcggtgtgaaactctggagagaaacagcaaat 1891
Db 3478 cagacataagaatttggaaataagccagcggtgtgaaactctggagagaaacagcaaat 3537
OY 1892 ccttagagatcagacatgtctgtgacagagagatgagagatcggttaagacatgagagag 1951
Db 3538 ccttagagatcagacatgtctgtgacagagagatgagagatcggttaagacatgagagag 3597
OY 1952 tgaatgagagagacatttcacacacatgtctcggagagatgacagctggcaaaatggaa 2011
Db 3598 tgaatgagagagacatttcacacacatgtctcggagagatgacagctggcaaaatggaa 3657
OY 2012 ccttgaacccctggagagatgaaatgtgaa----- 2039
Db 3658 ccttgaacccctggagagatgaaatgtgaa----- 3717
OY 2040 ----- 2039
Db 3718 atgaaacagctgtgaaagcatttcgaaagcaggaacccctgtagtcttattgtlaaga 3777
OY 2040 -----ggacccagctcagtgacagatcagagacagagacagagagctccat 2080
Db 3778 gcaattataaacagacaaacagacccacagctcagtgacagatcagagacagagagagctccat 3837
OY 2081 tgtcagatgtgccccacacccctccttcagagcttgcggaatgtgtgagatcaaacac 2140
Db 3838 tgtcagatgtgccccacacccctccttcagagcttgcggaatgtgtgagatcaaacac 3897
OY 2141 agtcatctgcagacaaaatctcacaagaatgtgcagacaaagagatgtgttctacagct 2200
Db 3898 agtcatctgcagacaaaatctcacaagaatgtgcagacaaagagatgtgttctacagct 3957
OY 2201 ggaaaaaatacagagagaggttatagtgaacccctaaacagcgagatcgatatgttaactcg 2260
Db 3958 ggaaaaaatacagagagaggttatagtgaacccctaaacagcgagatcgatatgttaactcg 4017
OY 2261 agaaaggtcagacagtggtctgggacctaaagctctgtcgtggaacaaagacagacagagatga 2320
Db 4018 agaaaggtcagacagtggtctgggacctaaagctctgtcgtggaacaaagacagacagagatga 4077
OY 2321 gtgtcttcatagttgaggatgtgatccaaaatggagctgcagagaaagatgttcgattgcaca 2380
Db 4078 gtgtcttcatagttgaggatgtgatccaaaatggagctgcagagaaagatgttcgattgcaca 4137
OY 2381 tttcagagatgagctctcagagatcaatgtgtcagattttatatagtgaagagatcaagaatg 2440
Db 4138 tttcagagatgagctctcagagatcaatgtgtcagattttatatagtgaagagatcaagaatg 4197
OY 2441 cctcatcaatcataaatgtgcccccttccttaaaagtgaataatattttatcgaataaag 2500
Db 4198 cctcatcaatcataaatgtgcccccttccttaaaagtgaataatattttatcgaataaag 4257
OY 2501 atgcagatgaatcagatggtcgatgtcctggaatgtcagtaagaaccttgcctctaact 2560
Db 4258 atgcagatgaatcagatggtcgatgtcctggaatgtcagtaagaaccttgcctctaact 4317
OY 2561 cagaaatctcacaataaagagacagagccaactgttactactctgagtcagctgtg 2620
Db 4318 cagaaatctcacaataaagagacagagccaactgttactactctgagtcagctgtg 4377
OY 2621 accctcagttcatttaaaaatgtgcaacatctgagcttcccaagatcagaggggttgg 2680
Db 4378 accctcagttcatttaaaaatgtgcaacatctgagcttcccaagatcagaggggttgg 4437
OY 2681 gtaattgtcagcgaaagatacaactcagttgagatcatataaagagcttaacagag 2740
Db 4438 gtaattgtcagcgaaagatacaactcagttgagatcatataaagagcttaacagag 4497
OY 2741 atgggttagcagcgagatgtgaagctcaaaagtctgaagatcaatactgtgtagatg 2800
Db 4498 atgggttagcagcgagatgtgaagctcaaaagtctgaagatcaatactgtgtagatg 4557
OY 2801 atgaaatgtgtgtgttaccctattgaaaggtttatagccttctgaagacagaaaga 2860
Db 4558 atgaaatgtgtgtgttaccctattgaaaggtttatagccttctgaagacagaaaga 4617
OY 2861 tgaagataaaatccatccatctggaatccagatcccaagctgtcttcacagag 2920
Db 4618 tgaagataaaatccatccatctggaatccagatcccaagctgtcttcacagag 4677
OY 2921 ctgttgacagcgagtggaagaaaaaagacagctccacagctctcgtatgtgtccacagctg 2980
Db 4678 ctgttgacagcgagtggaagaaaaaagacagctccacagctctcgtatgtgtccacagctg 4737
OY 2981 gtcocccagacccgagatccatccgaataacagatcaatccacacagcaattttg 3040
Db 4738 gtcocccagacccgagatccatccgaataacagatcaatccacacagcaattttg 4797
OY 3041 ctctgacatcgtcgaacccctgacacatctatccctgtgctggaacacacatcgagattcca 3100
Db 4798 ctctgacatcgtcgaacccctgacacatctatccctgtgctggaacacacatcgagattcca 4857
OY 3101 aagggcgaaacggctgtggtcgtgagatctgtggtgttcagacagctgtctgtgtgct 3160
Db 4858 aagggcgaaacggctgtggtcgtgagatctgtggtgttcagacagctgtctgtgtgct 4917
OY 3161 ttattatccatgaagtttatgaagagagagagacatgtgaagatcgtgagctg 3220
Db 4918 ttattatccatgaagtttatgaagagagagagacatgtgaagatcgtgagctg 4977
OY 3221 gagaatcagatcttagaggtgaaatgtgaatctgaaagagagacacatgtgaagacaa 3280
Db 4978 gagaatcagatcttagaggtgaaatgtgaatctgaaagagagacacatgtgaagacaa 5037
OY 3281 tcaatgtctcagacagacgacagagagagcgccctgacacctctacagagatgaagccc 3340
Db 5038 tcaatgtctcagacagacgacagagagagcgccctgacacctctacagagatgaagccc 5097
OY 3341 catcaaaagagagagagatgtgtgacacccctacatattgaaactgcagaaagagccggaa 3400
Db 5098 catcaaaagagagagagatgtgtgacacccctacatattgaaactgcagaaagagccggaa 5157
OY 3401 aagggctagaatgaatgtgtgtaaaagaaacgatactgagatgtgtgtcagaca 3460
Db 5158 aagggctagaatgaatgtgtgtaaaagaaacgatactgagatgtgtgtcagaca 5217
OY 3461 tttcacaagagagagatgtgcataatcccgatgaaagacagatccagagagacagatattat 3520
Db 5218 tttcacaagagagagatgtgcataatcccgatgaaagacagatccagagagacagatattat 5277
OY 3521 tgtgtgaatgggagagctgtcgtatgtcctcccaagagcggtgtgcctgtgttaagt 3580
Db 5278 tgtgtgaatgggagagctgtcgtatgtcctcccaagagcggtgtgcctgtgttaagt 5337
OY 3581 gtcccttagagcagataactctggaatgtgaaagatacaagctgtgtcatccatcag 3640
Db 5338 gtcccttagagcagataactctggaatgtgaaagatacaagctgtgtcatccatcag 5397
OY 3641 agagagagccatctcaaaacagacagtgagtgaaagcagctgtctcttcaacttc 3700
Db 5398 agagagagccatctcaaaacagacagtgagtgaaagcagctgtctcttcaacttc 5457
OY 3701 cactctctgatacagataactctgagctcagtaagatagctcaaaagaaatgtcattg 3760
Db 5458 cactctctgatacagataactctgagctcagtaagatagctcaaaagaaatgtcattg 5517

```

Db      6598  ccctcaattgacattcactctctctccaccaccaact 6637
|||||
RESULT# 4
ID      AAX29910
AC      AAX29910 standard; DNA; 6750 BP.
XX      AAX29910;
XX      06-JUL-1999 (first entry)
DE      Clone 38-2-1b encoding protein containing PDZ domain.
XX      PDZ domain; gene expression: human umbilical vascular endothelial cell;
XX      HUVEC; stimulation: tumour necrosis factor; TNF; protein binding;
XX      cell; proliferation disorder; cancer; ss.
OS      Homo sapiens.
XX      WO9907846-A1.
XX      18-FEB-1999.
XX      12-AUG-1998; 98WO-JP03603.
XX      19-JUN-1998; 98JP-0189944.
XX      12-AUG-1997; 97JP-0230356.
XX      (CHUG- ) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX      Funahashi S, Miyata S;
XX      WPI: 1999-167423/14.
XX      P-PSDB; AAY04733.
XX      Protein containing PDZ domain, whose expression is enhanced by TNF
XX      stimulation - plays an important role in protein/protein
XX      interactions and is used for screening for proteins for use in
XX      treatment of cell proliferation disorders such as cancer
XX      Example 7; Page 202-221; 240pp; Japanese.
XX      This sequence represents clone 38-2-1b which encodes a new protein.
XX      containing PDZ domains whose expression in human umbilical vascular
XX      endothelial cells (HUVEC) is enhanced by stimulation with tumour
XX      necrosis factor (TNF) alpha. The new protein is used to identify
XX      proteins which bind to it (particularly to the PDZ domains) and the
XX      CC genes encoding them, for use in the treatment of cell proliferation
XX      disorders such as cancer.
XX      Sequence 6750 BP; 2054 A; 1385 C; 1634 G; 1677 T; 0 other:
SO
Query Match 90.9%; Score 4436.2; DB 20; Length 6750;
Best Local Similarity 95.5%; Pred. No. 0;
Matches 4661; Conservative 0; Mismatches 8; Indels 210; Gaps 1

```


Dd	5352	atcccgatcgaaagactctgataccaaaggagacacagatatattgtgtgaatggggaagacgttc	5411
Qy	3542	gtaatgcctcccaagaagcggtgtgcggtttgtctaagtgtctccatgagcagatlaacct	3601
Dd	5412	gtaatgcctcccaagaagcggtgtgcggtttgtctaagtgtctccatgagcagatlaacct	5471
Qy	3602	tcggaagtgcggaagatacaaaagctgcgtccattccattcagaaggagggcgatctcaaacca	3661
Dd	5412	tcggaagtgcggaagatacaaaagctgcgtccattccattcagaaggagggcgatctcaaacca	5511
Qy	3662	gccagctcgagatcgaaagcgacgcctgtctctcttcaccttccacatctctctgataccagtaacat	3721
Dd	5532	gccagctcgagatcgaaagcgacgcctgtctctctcttcaccttccacatctctctgataccagtaacat	5591
Qy	3722	ctgagctacatcgtgaaagaagtacgtccaagaagaagatgcatctgtgacatctgaaataacagggtatcaa	3781
Dd	5592	ctgagctacatcgtgaaagaagtacgtccaagaagaagatgcatctgtgacatctgaaataacagggtatcaa	5651
Qy	3782	gaacacgtcgaatactgaataaaagcgccctacatgacacatctggaaatcgagatctgcgtctgaagag	3841
Dd	5652	gaacacgtcgaatactgaataaaagcgccctacatgacacatctggaaatcgagatctgcgtctgaagag	5711
Qy	3842	taggcagccacacatctgtgtgatactgataatctatctgcaatgtgatgcccaacacgtgagttg	3901
Dd	5712	taggcagccacacatctgtgtgatactgataatctatctgcaatgtgatgcccaacacacgtgagttg	5771
Qy	3902	cagcgcacaaacccaacaaactctagaagttctggggaatggatcttgcacacatctgtgtgcacatacca	3961
Dd	5772	cagcgcacaaacccaacaaactctagaagttctggggaatggatcttgcacacatctgtgtgcacatacca	5831
Qy	3962	ctgagtgagatgatactcaacacccaacagcagtttaaacctacatgaaaaatgcatctggtccatctg	4021
Dd	5832	ctgagtgagatgatactcaacacccaacagcagtttaaacctacatgaaaaatgcatctggtccatctg	5891
Qy	4022	aaatgacaggtgtgtgtgcgtggagggagacgttgagtggtgtccaaagtcatcatacaaggagccgtg	4081
Dd	5892	aaatgacaggtgtgtgtgcgtggagggagacgttgagtggtgtccaaagtcatcatacaaggagccgtg	5951
Qy	4082	caaatctccagactctctcttcaactggtgcgtgcagcgtcaaacacagatatacttcagagatgatactag	4141
Dd	5952	caaatctccagactctctcttcaactggtgcgtgcagcgtcaaacacagatatacttcagagatgatactag	6011
Qy	4142	gaactctccaaatgtgaagtctcatctaacctctagagcgagagacccaagatggtcttaagctccagta	4201
Dd	6012	gaactctccaaatgtgaagtctcatctaacctctagagcgagagacccaagatggtcttaagctccagta	6071
Qy	4202	tagtctgagtgagatactgagagacccctacatgagagacttaaccattatctgttaaaacagatcttg	4261
Dd	6072	tagtctgagtgagatactgagagacccctacatgagagacttaaccattatctgtgttaaaacagatcttg	6131
Qy	4262	caaaagggagcagcctctgaaagacgagacgtctgaaaaaggggcgatccagatcatctgtctca	4321
Dd	6132	caaaagggagcagcctctgaaagacgagacgtctgaaaaaggggcgatccagatcatctgtctca	6191
Qy	4322	atgggcacaaagctctagaaagggtgtaccacatgaaagaagctgttgcacatcccttaaacggagaca	4381
Dd	6192	atgggcacaaagctctagaaagggtgtaccacatgaaagaagctgttgcacatcccttaaacggagaca	6251
Qy	4382	aaggagacgttcaactctgaatgtctctctctctgaatctgtgcgtgcagagaattgtgaaccaaccca	4441
Dd	6252	aaggagacgttcaactctgaatgtctctctctctgaatctgtgcgtgcagagaattgtgaaccaaccca	6311
Qy	4442	ccactagctcaactctctactactgttaagaagaatgacatcggtctccgtgacaattttatagctgtg	4501
Dd	6312	ccactagctcaactctctactactgttaagaagaatgacatcggtctccgtgacaattttatagctgtg	6371
Qy	4502	ttcagccggggtcttcaaaactgtaggggggaaataaaccttgaattctctttccatct	4561
Dd	6372	ttcagccggggtcttcaaaactgtaggggggaaataaaccttgaattctctttccatct	6431
Qy	4562	agaaatgcttctccttactgcaaacctaacatcatcttctctctctctgtcatctttgtga	4621
Dd	6432	agaaatgcttctccttactgcaaacctaacatcatcttctctctctctgtcatctttgtga	6491

QY	4622	acttaaaagaagaagaataatttgcgttagcggaatccgcgttttaatttgcgtgagatccaa	4681
Db	6492	acttaaaagaagaagaataatttgcgttagcggaatccgcgttttaatttgcgtgagatccaa	6551
QY	4682	tgtttgcgttagcgcattgcggcgaagaattatataatgcctaagctgcgttagtataaagaaga	4741
Db	6552	tgtttgcgttagcgcattgcggcgaagaattatataatgcctaagctgcgttagtataaagaaga	6611
QY	4742	taattctaaagctcaacccaagaagaatggcttaagtaagttagagatgaaataatgaaatat	4801
Db	6612	taattctaaagctcaacccaagaagaatggcttaagtaagttagagatgaaataatgaaatat	6671
QY	4802	aaaaataaagaagaanaatctctcggggagtttaaaaaaaatgcctcaatttgcgaatctact	4861
Db	6672	aaaaataaagaagaanaatctctcggggagtttaaaaaaaatgcctcaatttgcgaatctact	6731
QY	4862	cctctcccccaccaccaact	4880
Db	6732	cctctcccccaccaccaact	6750

RESULT	5
AAA47424	
ID	AAA47424 standard; DNA; 3188 BP.

AC AAA47424;

DT 20-OCT-2000 (first entry)

DE Sequence encoding human neuron-associated protein

KM Neuroan associated protein; NEUAP; neurologic disorder; epilepsy;
KM ischemic cerebrovascular disease; stroke; cerebral neoplasm;
KM Alzheimer's disease; Pick's disease; Huntington's disease;
KM dementia; Parkinson's disease; demyelinating disease; meningitis;
KM prion disease; kuru; Creutzfeldt-Jakob disease; neurofibrinomas;
KM cerebral palsy; muscular dystrophy; central nervous system; CNS;
KM peripheral nervous system; PNS; myopathy; schizophrenia;
KM actinic keratosis; arteriosclerosis; atherosclerosis; buritis;
KM cirrhosis; hepatitis; mixed connective tissue disease; MCTD;
KM myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;
KM autoimmune disease; inflammation; acquired immunodeficiency syndrome;
KM AIDS; Addison's disease; adult respiratory distress syndrome;
KM allergy; ankylosing spondylitis; amyloidosis; anaemia; asthma;
KM Werner syndrome; trauma; human; ds.
XS Homo sapiens.

Homo sapiens

FH	Key	Location/Qualifiers
FT	CDS	
		29..2599

/product= Neuron associated protein

PN WO200034477-A2

PD 15-JUN-2000.

PF 10-DEC-1999;

PR 11-DEC-1998;

PR 09-FEB-1999; 99US-0119365.

XX
XX
XX (TWO) TWOED DUTY TWO

XX
XX
E
B
W
F
P
S
T
L
J
A
H
I

PI LU DAM, Azimzai Y;

DR WPI; 2000-423423/36

XX

PT New human neuron-associated proteins and polynucleotides encoding them
PT useful for diagnosis, treatment and prevention of cell proliferative
PT disorders including cancer, neuronal and neurological disorders
XX
PS Claim 9; Page 131-132; 145pp; English.
XX

PS Claim 9; Page 131-132; 145pp; English.

Human neuron-associated proteins (NEUAP) can be used for treating or preventing a disorder associated with decreased expression or activity of NEUAP. Antagonists of NEUAP are useful for treating or preventing disorder associated with increased expression or activity of NEUAP. NEUAP or their fragments or derivatives are useful for treating neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia and Parkinson's disease. NEUAPs are also useful for treating other demyelinating diseases, bacterial and viral meningitis, prion diseases including kuru, Creutzfeldt-Jakob disease, nutritional and metabolic diseases of the nervous system, neurofibromatosis, other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, inherited, metabolic, endocrine, and toxic myopathies, mental disorders including mood, anxiety and schizophrenic disorders, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, and an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemias, asthma, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal parasitic, protozoal, and helminthic infections, and trauma. This sequence was given the Incyte ID no. 3201/53CBI.

SQ Sequence 3188 BP; 990 A; 644 C; 771 G; 783 T; 0 other;

Query match	57.98; Score 2825; DB 21; Length 3188;
-------------	--

Matches 2831; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2040 ggcaccagtcagtcagagtcagagccagagaagctccattgtgcagtggtgcccccacc 2099

Db 226 ggcacccagtcagtcagagcagagaagctccattgtgcagtggtgccccacc 285

QY 2100 ccctcctcagcctttgccgaaatgggtagtgatcacacacagtcattctgcaagcaaat 2159

Db 286 ccctccttcagcctttgcgaaatggtagtcgatacacacacagtcattctgcagaacaaat 345

QY 2160 ctcaagatgtgacaaagagatgtagcttggttacagctggaataatcagagagcg 2219

Db 346 ctcaacagatgtggaacaaagagatgagtttggttacagctggaataatcagagagcg 405

2220 ttatggaacctaacagcgagctgcataatgaactggagaaggtcatagtgtt 2279

Db 406 ttatggaaccctaacagcgagctgcataatgtgaactggagaaggtcatagtggtt 465

QY 2280 gggcctaagtcctgctggacaagaccgatccagatgagtgctctcatagtgggat 2339

Db 466 gggcctaagtcctgtcggacaagaccgatccagatgtagtgtcctcatagtgggat 525

2340 tgatccaatgtagctgcaggaaagatgtgcgattgcaaatgtagctctaga 2399

Db 526 tgatccaatgagctgcaggaagaatgtgcgattgcaaatgcaagatgagcttctaga 585

2400 gaccaatggtcagattttatatggaagaatcatcagaatgcctcatcaatcataatg 2459

Db 586 gatcaatggtcagattttatatgtgaagaagtcatacagaatgcctcatcaatcataaattg 645

QY 2460 tggcccttctaagtgaataattttatcagaataaagatgcagtgaatcagatggc 2519

|||||
Db 646 tgcctcttaagtgaaataattttatctcgaataaagatgcgtgaatcaagtggtc 705
Qy 2520 cgtatgtcttgaaatgcsagtgaaccttgccttctaactgaataatctcaataa 2579
Db 706 cgtatgtcttgaaatgcsagtgaaccttgccttctaactgaataatctcaataa 765
Qy 2580 ggaagaagagccaaactgttaactctctgatgagcttgaggcctaagtcttaataa 2639
Db 766 ggaagaagagccaaactgttaactctctgatgagcttgaggcctaagtcttaataa 825
Qy 2640 tgtgcaacactctgagactctcccaagagatcaggggtcttggtatctgcatcagcga 2699
Db 826 tgtgcaacactctgagactctcccaagagatcaggggtcttggtatctgcatcagcga 885
Qy 2700 agatcacactcagtgagatcaataaagacttaacagagcatggtgtatgacagcagga 2759
Db 886 agatcacactcagtgagatcaataaagacttaacagagcatggtgtatgacagcagga 945
Qy 2760 tggagcactcaaaagtctggagatcagatactagctgtgagatgataatgtgtgtta 2819
Db 946 tggagcactcaaaagtctggagatcagatactagctgtgagatgataatgtgtgtta 1005
Qy 2820 cccatctgaaaagtcttatagcctctctgaaagacagagatgacagtaaaacttaacat 2879
Db 1006 cccatctgaaaagtcttatagcctctctgaaagacagagatgacagtaaaacttaacat 1065
Qy 2880 ccaatctgagaatccagatccagagctgtctctcaagagctgtgtgacagctgagga 2939
Db 1066 ccaatctgagaatccagatccagagctgtctctcaagagctgtgtgacagctgagga 1125
Qy 2940 aaaaaagaaagctcccaagctctgtatgtgtcccaagctgtgtgtcccaaaacagagtc 2999
Db 1126 aaaaaagaaagctcccaagctctgtatgtgtcccaagctgtgtgtcccaaaacagagtc 1185
Qy 3000 catcctgaaatacaagagatcatcaacaacacagcaattttgtctctgactctgcaactg 3059
Db 1186 catcctgaaatacaagagatcatcaacaacacagcaattttgtctctgactctgcaactg 1245
Qy 3060 ccccatatccctgtgctgcaaaaacacacatcgagatttccaaaggcgaaacaggtcg 3119
Db 1246 ccccatatccctgtgctgcaaaaacacacatcgagatttccaaaggcgaaacaggtcg 1305
Qy 3120 cctgagacatctgtggggtctcagaacagctgtcgggtgccttattatctcaataagtta 3179
Db 1306 cctgagacatctgtggggtctcagaacagctgtcgggtgccttattatctcaataagtta 1365
Qy 3180 tgaagaagagcagcatgttaaaagtgaagaactctggctgagatcaaga tcttaaggtc 3239
Db 1366 tgaagaagagcagcatgttaaaagtgaagaactctggctgagatcaaga tcttaaggtc 1425
Qy 3240 gaattggaattgactctgaggaagccacacatagtatgaagaatcaatgtcctgagacagc 3299
Db 1426 gaattggaattgactctgaggaagccacacatagtatgaagaatcaatgtcctgagacagc 1485
Qy 3300 gcaacagagagtgctgctgagcatcttaacagatgagggcccttaaaagagagagagc 3359
Db 1486 gcaacagagagtgctgctgagcatcttaacagatgagggcccttaaaagagagagagc 1545
Qy 3360 gcttgacaacctcaacttaagctgcaagaagccgggaaaaagcccttaagatcaat 3419
Db 1546 gcttgacaacctcaacttaagctgcaagaagccgggaaaaagcccttaagatcaat 1605
Qy 3420 tgttggtaaaagaaacagatactggaatattgtgtcagacatgtcctaaagaggaattgc 3479
Db 1606 tgttggtaaaagaaacagatactggaatattgtgtcagacatgtcctaaagaggaattgc 1665
Qy 3480 agatcccgatggaagactgttccagagagacagagataattgtgtatgtgggaagagcgt 3539
Db 1666 agatcccgatggaagactgttccagagagacagagataattgtgtatgtgggaagagcgt 1725
Qy 3540 tcttaatgtctcccaagaaacggtgtgcgcttctgctaaatgtctcccttaagccagtaac 3599
|||||

Db 1726 tcttaatgtctcccaagaaacggtgtgcgcttctgctaaagtgtctcccttaagccagtaac 1785
Qy 3600 ctttggaagtctggaagaatacaaaagctgtgtccattcttaagagagagccattcaaac 3659
Db 1786 ctttggaagtctggaagaatacaaaagctgtgtccattcttaagagagagccattcaaac 1845
Qy 3660 cagccagctggaatgaaagagcagctgtctctcttcaactcttccactctcggatccagtaac 3719
Db 1846 cagccagctggaatgaaagagcagctgtctctcttcaactcttccactctcggatccagtaac 1905
Qy 3720 atctgagctactggaagaagtaagctcaaaagaaatgcatctgacatctgaaatacaaggaatt 3779
Db 1906 atctgagctactggaagaagtaagctcaaaagaaatgcatctgacatctgaaatacaaggaatt 1965
Qy 3780 aagaacagctcgaataatgaaaaagggcctgtgactcaactcaggggaatcagatcgttgagg 3839
Db 1966 aagaacagctcgaataatgaaaaagggcctgtgactcaactcaggggaatcagatcgttgagg 2025
Qy 3840 agtaagcagcccaactgtgtgtgtgtctatattatctgcaatgtgacaccaactgtgagtc 3899
Db 2026 agtaagcagcccaactgtgtgtgtgtgtctatattatctgcaatgtgacaccaactgtgagtc 2085
Qy 3900 tgcagcagcagcccaaaactcaagagctggtggaatgagatgtgcaacatctgtgacatc 3959
Db 2086 tgcagcagcagcccaaaactcaagagctggtggaatgagatgtgcaacatctgtgacatc 2145
Qy 3960 cactggggcagtactcacaacccaagagtttaacctcagaaatgcatctgtgcctcat 4019
Db 2146 cactgggggagatgactcacaacccaagagtttaacctcagaaatgcatctgtgcctcat 2205
Qy 4020 tgaatgcagagctgtgtctggaagaagacgtggaatgtgtgtcaacagctcatcatcagaagc 4079
Db 2206 tgaatgcagagctgtgtctggaagaagacgtggaatgtgtgtcaacagctcatcatcagaagc 2265
Qy 4080 tgcagatccagatctctcttcaactgtgtgacgtgcaacagatattctcaggaatgtatt 4139
Db 2266 tgcagatccagatctctcttcaactgtgtgacgtgcaacagatattctcaggaatgtatt 2325
Qy 4140 aggaactccctcaatgttaagtcttatcaactaagagcgaagccagatgtgactgaactcag 4199
Db 2326 aggaactccctcaatgttaagtcttatcaactaagagcgaagccagatgtgactgaactcag 2385
Qy 4200 tatagttgagagataatgagccctcatgtagagacttaacattatgttaaaacagtgct 4259
Db 2386 tatagttgagagataatgagccctcatgtagagacttaacattatgttaaaacagtgct 2445
Qy 4260 tgcnaagggagcagcctctgaaagcaggaactgtgaaagggcgatcagatcaatctgtct 4319
Db 2446 tgcnaagggagcagcctctgaaagcaggaactgtgaaagggcgatcagatcaatctgtct 2505
Qy 4320 caatgggcagagctctgaagagatcaacccaatgaagaagctgtgtgcatccttaaacgagc 4379
Db 2506 caatgggcagagctctgaagagatcaacccaatgaagaagctgtgtgcatccttaaacgagc 2565
Qy 4380 aaaaagcagctcaactctgtgtctctcttgaatgtgctgcagagatgaaacaaacc 4439
Db 2566 aaaaagcagctcaactctgtgtctctcttgaatgtgctgcagagatgaaacaaacc 2625
Qy 4440 aaccctcagctcaactcctcaatgtaagaagatgacatgtgtcctgacaattttatgtctg 4499
Db 2626 aaccctcagctcaactcctcaatgtaagaagatgacatgtgtcctgacaattttatgtctg 2685
Qy 4500 tgttcagcgggtcttcaaaactgttgagggggaataaaccataagtcttcttctcat 4559
Db 2686 tgttcagcgggtcttcaaaactgttgagggggaataaaccataagtcttcttctcat 2745
Qy 4560 ctagaatgtcttctcactcagacaacccaatcatttcttcttctctgtgactgtgt 4619
Db 2746 ctagaatgtcttctcactcagacaacccaatcatttcttcttctctgtgactgtgt 2805
Qy 4620 gaacttaagaagaagataattgtgtgaagtgaaatcgttttattatgttgagagatact 4679
Db 2806 gaacttaagaagaagataattgtgtgaagtgaaatcgttttattatgttgagagatact 2865
|||||

OY 4680 aatgttttgatgacatgagcaagaattattacatgctaagctgtgagtaagaaga 4739
|||||
DB 2866 aatgttttgatgacatgagcaagaattattacatgctaagctgtgagtaagaaga 2925
OY 4740 gataattctaaagctaacccaagaataatgcttcagtaagttagagtaaaaaatgaaat 4799
|||||
DB 2926 gataattctaaagctaacccaagaataatgcttcagtaagttagagtaaaaaatgaaat 2985
OY 4800 ataataaagaagaataatctggggagttttaaanaaaatgctcaatttggcaatctac 4859
|||||
DB 2966 ataataaagaagaataatctggggagttttaaanaaaatgctcaatttggcaatctac 3045
OY 4860 ctccctcccccaccccaact 4880
|||||
DB 3046 ctccctcccccaccccaact 3066

RESULT 6
AAK29960
ID AAK29960 standard; DNA; 2819 BP.
XX
AC AAK29960;
DT 06-JUL-1999 (first entry)
XX
DE PDZ domain-containing protein gene clone 32-8-1/5R3.
XX
KW PDZ domain: gene expression; human umbilical vascular endothelial cell;
KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding; PCR;
KW cell; proliferation disorder; cancer; primer; amplification; ss.
XX
OS Homo sapiens.
XX
PN MO9907846-A1.
PD 18-FEB-1999.
XX
PF 12-AUG-1998; 98WO-JP03603.
XX
PR 19-JUN-1998; 98JP-0189944.
PR 12-AUG-1997; 97JP-0230356.
XX
PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
PI Funahashi S, Miyata S;
XX
DR WPI: 1999-167423/14.
DR P-PSDB: AAY04741.
XX
PT Protein containing PDZ domain, whose expression is enhanced by TNF
PT stimulation - plays an important role in protein/protein
PT interactions and is used for screening for proteins for use in
PT treatment of cell proliferation disorders such as cancer
XX
PS Example 2; Page 127-134; 240pp; Japanese.
XX
CC This sequence represents clone 32-8-1/5R3 which encodes a new protein
CC containing PDZ domains whose expression in human umbilical vascular
CC endothelial cells (HUVEC) is enhanced by stimulation with tumour necrosis
CC factor (TNF) alpha. The new protein is used to identify proteins which
CC bind to it (particularly to the PDZ domains) and the genes encoding them,
CC for use in the treatment of cell proliferation disorders such as cancer.
XX
SQ Sequence 2819 BP; 865 A; 583 C; 678 G; 693 T; 0 other;

Query Match 57.0%; Score 2781.8; DB 20; Length 2819;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2783; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2096 caccctccctcagccttgcgcgaatgagtgatgacacacagtcattctgcaaga 2155
|||||

DB 17 caccctccctcagccttgcgcgaatgagtgatgacacacagtcattctgcaaga 76
OY 2156 aatctcaagaatgtggacaagaagagtgattgtgttacaagctgggaaaaatcacag 2215
|||||
DB 77 aatctcaagaatgtggacaagaagagtgattgtgttacaagctgggaaaaatcacag 136
OY 2216 agcgttatgaaaccccaacaaagcgagcgcacatgatgaaactggagaaagtcataagtg 2275
|||||
DB 137 agcgttatgaaaccccaacaaagcgagcgcacatgatgaaactggagaaagtcataagtg 196
OY 2276 gtttggccttaagcttctctgtgggaacaaagccatcagaatgagtgcttctcatagtg 2335
|||||
DB 197 gtttggccttaagcttctctgtgggaacaaagccatcagaatgagtgcttctcatagtg 256
OY 2336 ggaattgaccaaagtgaagctcagaagaagaatgagtcgaattcgaattgagatgagctc 2395
|||||
DB 257 ggaattgaccaaagtgaagctcagaagaagaatgagtcgaattcgaattgagatgagctc 316
OY 2396 tagagatcaatggtcagatttataatggaagaagtcacatgaatgctccatcaatcata 2455
|||||
DB 317 tagagatcaatggtcagatttataatggaagaagtcacatgaatgctccatcaatcata 376
OY 2456 aatgtgccctctctaaagtgaataaattttatcagaataaagaatgacatgacatcaga 2515
|||||
DB 377 aatgtgccctctctaaagtgaataaattttatcagaataaagaatgacatgacatcaga 436
OY 2516 tggccgtatgtctcgtgaatcagtagaaccttgccttcaacctcagaataatctcaca 2575
|||||
DB 437 tggccgtatgtctcgtgaatcagtagaaccttgccttcaacctcagaataatctcaca 496
OY 2576 ataagagaacagagcacaactgttactactctgatgacagctgtgagcctcagttcata 2635
|||||
DB 497 ataagagaacagagcacaactgttactactctgatgacagctgtgagcctcagttcata 556
OY 2636 aaaatgtcaacatctgagagcttcccaagagatcaaggagggttgggtatctcaccagcg 2695
|||||
DB 557 aaaatgtcaacatctgagagcttcccaagagatcaaggagggttgggtatctcaccagcg 616
OY 2696 aagaagatacactcagatgagatcatcataaagaagcttaacagagatgggttagcagcca 2755
|||||
DB 617 aagaagatacactcagatgagatcatcataaagaagcttaacagagatgggttagcagcca 676
OY 2756 cggatgagcagactcaaaatcgagagatcagatacagctgtgtatgagatgaaatgtgtgtg 2815
|||||
DB 677 cggatgagcagactcaaaatcgagagatcagatacagctgtgtatgagatgaaatgtgtgtg 736
OY 2816 gttaccctatgaaagaattttatagccttctggaagaacagcaaaagtgtgacataaacta 2875
|||||
DB 737 gttaccctatgaaagaattttatagccttctggaagaacagcaaaagtgtgacataaacta 796
OY 2876 ccatcatgctgagatccagatccatccagagctgttccctcagcagctgtgtgacagcagtg 2935
|||||
DB 797 ccatcatgctgagatccagatccatccagagctgttccctcagcagctgtgtgacagcagtg 856
OY 2936 gagaanaaagaacagctcccaagctctctgtagtgttcccaagctgtgctccccaagaccg 2995
|||||
DB 857 gagaanaaagaacagctcccaagctctctgtagtgttcccaagctgtgctccccaagaccg 916
OY 2996 agtcatcagaataacagagcagatcatcaaacacagcaaatgttgccttcatctcagtgaa 3055
|||||
DB 917 agtcatcagaataacagagcagatcatcaaacacagcaaatgttgccttcatctcagtgaa 976
OY 3056 cctggcccatatccctctgctgtcgaaacaacatcagaatcttccaaagggcgaaacagggc 3115
|||||
DB 977 cctggcccatatccctctgctgtcgaaacaacatcagaatcttccaaagggcgaaacagggc 1036
OY 3116 tgggcctgagcagctgtgggggttcaagacagcgtgtgtgtgctttatcatatgaag 3175
|||||
DB 1037 tgggcctgagcagctgtgggggttcaagacagcgtgtgtgtgctttatcatatgaag 1096
OY 3176 ttatgaagaagagcagacatgataagaatggaagccttgggcctggaataagatcttag 3235
|||||
DB 1097 ttatgaagaagagcagacatgataagaatggaagccttgggcctggaataagatcttag 1156

QY 3236 aggtgaatggaattgagtgaggaaagccacacatgatgaagaatcaatgctctagac 3295
 |||||||
 Db 1157 aggtgaatggaattgacctgaggaagccacacatgatgaagaatcaatgctctagac 1216
 QY 3296 agagcgacagagaagtgccgctgacactctacagagatgagggccacatacaagaagagag 3355
 |||||||
 Db 1217 agagcgacagagaagtgccgctgacactctacagagatgagggccacatacaagaagagag 1276
 QY 3356 aagtgctgacacccctcaactattgagctgcagaagaagccgggaaaaagcgctagattaa 3415
 |||||||
 Db 1277 aagtgctgacacccctcaactattgagctgcagaagaagccgggaaaaagcgctagattaa 1336
 QY 3416 gttcttggtgtaaaagaagaatgactgagatatttgtragaacattgttcaaaaggaggaa 3475
 |||||||
 Db 1337 gttcttggtgtaaaagaagaatgactgagatatttgtragaacattgttcaaaaggaggaa 1396
 QY 3476 ttgcagatcccgatggaagaatgatccagaggagaacagatatattgtgtgaatggggaag 3535
 |||||||
 Db 1397 ttgcagatcccgatggaagaatgatccagaggagaacagatatattgtgtgaatggggaag 1456
 QY 3536 aagttcgttaattgctccccaagaagcggtgctgagcttgcctaaagtgtccctagagcaag 3595
 |||||||
 Db 1457 aagttcgttaattgctccccaagaagcggtgctgagcttgcctaaagtgtccctagagcaag 1516
 QY 3596 taaccttggaagtgtggaagaatacaaaagctggtccattccattcagaagagagccatctc 3655
 |||||||
 Db 1517 taaccttggaagtgtggaagaatacaaaagctggtccattccattcagaagagagccatctc 1576
 QY 3656 aaacccagccagtgtagtggaagcagcctgtcctcttcaatttccactctcggatcca 3715
 |||||||
 Db 1577 aaacccagccagtgtagtggaagcagcctgtcctcttcaatttccactctcggatcca 1636
 QY 3716 gtacatctggaatgactggaagaatgactcaaaagaagaatgcatgtgcatctgaaatacag 3775
 |||||||
 Db 1637 gtacatctggaatgactggaagaatgactcaaaagaagaatgcatgtgcatctgaaatacag 1696
 QY 3776 gattaagaagaatgcgaataaagaagggccctctgactcactcgggaatgacatcgctg 3835
 |||||||
 Db 1697 gattaagaagaatgcgaataaagaagggccctctgactcactcgggaatgacatcgctg 1756
 QY 3836 gaaggaatgagcagccacactggtgcatgtgctctatttattgtcaatgatgaccacactg 3895
 |||||||
 Db 1757 gaaggaatgagcagccacactggtgcatgtgctctatttattgtcaatgatgaccacactg 1816
 QY 3896 gaagtcgacagacagaccacaaactcaagagtggggaatgattgtcacatctgtgca 3955
 |||||||
 Db 1817 gaagtcgacagacagaccacaaactcaagagtggggaatgattgtcacatctgtgca 1876
 QY 3956 catccactgaggcagatgactcacacccaagaatctaacactgaanaaatagcatctggtc 4015
 |||||||
 Db 1877 catccactgaggcagatgactcacacccaagaatctaacactgaanaaatagcatctggtc 1936
 QY 4016 ccaattgaatgagcagtggtgctggaagaacgtgagtggtgcacagagtcacatcag 4075
 |||||||
 Db 1937 ccaattgaatgagcagtggtgctggaagaacgtgagtggtgcacagagtcacatcag 1996
 QY 4076 agcctgcaagttcagctcttcttcttcactggtgcagtcacacagataattcagatg 4135
 |||||||
 Db 1997 agcctgcaagttcagctcttcttcttcactggtgcagtcacacagataattcagatg 2056
 QY 4136 atttaagacctctcgaatgaatgatactatactatagagagagacagatggttggct 4195
 |||||||
 Db 2057 atttaagacctctcgaatgaatgatactatactatagagagagacagatggttggct 2116
 QY 4196 tcaagtaagtgtggaagataagcagccctcatatgagagactaacatattatgttaaaacag 4255
 |||||||
 Db 2117 tcaagtaagtgtggaagataagcagccctcatatgagagactaacatattatgttaaaacag 2176
 QY 4256 tgtttgcaaaaggagcagcctctgaaagcagagctctgaaagggcgatcagatcatg 4315
 |||||||
 Db 2177 tgtttgcaaaaggagcagcctctgaaagcagagctctgaaagggcgatcagatcatg 2236

QY 4316 ctgtcaatggcagagctctagaagaagatcacccaatgaagaagctgttgcacatccttaac 4375
 |||||||
 Db 2237 ctgtcaatggcagagctctagaagaagatcacccaatgaagaagctgttgcacatccttaac 2296
 QY 4376 ggaacaaaaggcaactgtcacttctgattgtctctctctgtaatttggctgcagaaatgaaaca 4435
 |||||||
 Db 2297 ggaacaaaaggcaactgtcacttctgattgtctctctctgtaatttggctgcagaaatgaaaca 2256
 QY 4436 acccaacccctagctcactctcactgtaagaagaatgacatgctgtccatgaaattttat 4495
 |||||||
 Db 2357 acccaacccctagctcactctcactgtaagaagaatgacatgctgtccatgaaattttat 2416
 QY 4496 gctgttcaagccgggtcttccaactgtaggggggaataaacacttaagtttctttc 4555
 |||||||
 Db 2417 gctgttcaagccgggtcttccaactgtaggggggaataaacacttaagtttctttc 2476
 QY 4556 tcaatagaatgtcttcttcttactgacaacactacacatcttcttcttcttcttcttcttct 4615
 |||||||
 Db 2477 tcaatagaatgtcttcttcttactgacaacactacacatcttcttcttcttcttcttcttct 2536
 QY 4616 ttgtgaacttaagaagaagaatattgtgtgagtgaaatcctgtttattttattgtggaat 4675
 |||||||
 Db 2537 ttgtgaacttaagaagaagaatattgtgtgagtgaaatcctgtttattttattgtggaat 2596
 QY 4676 atctaagtgtttgtgtaacatgagcaagaattattacatgcaagctgtgtgtaaa 4735
 |||||||
 Db 2597 atctaagtgtttgtgtaacatgagcaagaattattacatgcaagctgtgtgtaaa 2656
 QY 4736 gaaagtaattcttaaaagctaaccaagaagaatgtcttcagtaagttagaatgaagaatga 4795
 |||||||
 Db 2657 gaaagtaattcttaaaagctaaccaagaagaatgtcttcagtaagttagaatgaagaatga 2716
 QY 4796 aaatataaataaagaagaagaatctcggggagtttaaaaaaatgctcctaattggcaat 4855
 |||||||
 Db 2717 aaatataaataaagaagaagaatctcggggagtttaaaaaaatgctcctaattggcaat 2776
 QY 4856 ctaccctctctcccccacccaact 4880
 |||||||
 Db 2777 ctaccctctctcccccacccaact 2801

RESULT 7
 ID AAX29911 standard; DNA; 6666 BP.
 AC AAX29911:
 XX
 DT 06-JUL-1999 (first entry)
 XX
 DE Clone 38-2-1c encoding protein containing PDZ domain.
 KW PDZ domain; gene expression; human umbilical vascular endothelial cell;
 KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding;
 KW cell; proliferation disorder; cancer; ss.
 OS Homo sapiens.
 PN WO907846-A1.
 PD 18-FEB-1999.
 PF 12-AUG-1998; 98MO-JP03603.
 PR 19-JUN-1998; 98JP-0189944.
 PR 12-AUG-1997; 97JP-0230356.
 PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
 PI Funahashi S, Miyata S;
 XX
 XX MPI; 1999-167423/14.
 DR P-PSDB; AAY04734.
 XX


```

|||||
Db 601 aaactcaagcggagctcatctctaaagagcttaacagcagcaggggtagcacaccagcagctgga 660
Oy 2764 cgactcaaaagtcggagatacagatactgctgtatagatgataatgctgtgttaccct 2823
Db 661 cgactcaaaagtcggagatacagatactgctgtatagatgataatgctgtgttaccct 720
Oy 2824 attgaaaagtattattagctctctgaaagacagcaaaagtacacgttaaaacttaccatcct 2883
Db 721 attgaaaagttattattagctctctgaaagacagcaaaagtacacgttaaaacttaccatcct 780
Oy 2884 gcttgaaacccaagatcccaagcgtgtctctcaagcagcgtgtgacagccagctgagaanaa 2943
Db 781 gcttgaaacccaagatcccaagcgtgtctctcaagcagcgtgtgacagccagctgagaanaa 840
Oy 2944 aagaacagctcccaagctctgtatgctccacagctgtgctccccaagccagagatccatc 3003
Db 841 aagaacagctcccaagctctgtatgctccacagctgtgctccccaagccagagatccatc 900
Oy 3004 cgaataacaaagacagatcacaacacacagcaattttgtctctgatactcgaactgcccc 3063
Db 901 cgaataacaaagacagatcacaacacacagcaattttgtctctgatactcgaactgcccc 960
Oy 3064 attatccctgctgctgcaaaacacacacagatctccaagggcgagacagggctggcctg 3123
Db 961 attatccctgctgctgcaaaacacacacagatctccaagggcgagacagggctggcctg 1020
Oy 3124 agcatcgtgtgggggtcagacacagctgtcgtgcttattatccatgaagttaagaa 3183
Db 1021 agcatcgtgtgggggtcagacacagctgtcgtgcttattatccatgaagttaagaa 1080
Oy 3184 gaagagcagcagcagctgtaaaagtgaagactcgtgctggagatcagatcttgaagggat 3243
Db 1081 gaagagcagcagcagctgtaaaagtgaagactcgtgctggagatcagatcttgaagggat 1140
Oy 3244 ggaattgacttgaggaagacccacacatgtaagacatcaatgtccctgagacagacgcca 3303
Db 1141 ggaattgacttgaggaagacccacacatgtaagacatcaatgtccctgagacagacgcca 1200
Oy 3304 cagagagctgcgctgacactctcacagagatgagggcccatacaagaagaggaagctgt 3363
Db 1201 cagagagctgcgctgacactctcacagagatgagggcccatacaagaagaggaagctgt 1260
Oy 3364 gacacccctactatgtagctgcagaaagcgggaaagggccctaagattatgtt 3423
Db 1261 gacacccctactatgtagctgcagaaagcgggaaagggccctaagattatgtt 1320
Oy 3424 ggtlaaaagaaacgatacttgagatctgtctcagacatgttcaaaagaggaattgcagat 3483
Db 1321 ggtlaaaagaaacgatacttgagatctgtctcagacatgttcaaaagaggaattgcagat 1380
Oy 3484 ccgagatgaaagactgatacagaggaagacagataattgtgtgaatgggaaagcgttcgt 3543
Db 1381 gccgagatgaaagactgatacagaggaagacagataattgtgtgaatgggaaagcgttcgt 1440
Oy 3544 aatgcctcccaagaagcgggttcgctgtgtcctaa 3578
Db 1441 aatgcctcccaagaagcgggttcgctgtgtcctaa 1475

```

```

KW inflammatory disease; neoplasia; metabolic disorder; testis; placenta;
KW brain; T cell; spleen; lung; heart; rhabdomyosarcoma; endocrine system;
KW endocrinopathy; endocrine polyglandular syndrome; endocrinoma; sepsis;
KW endocrine ophthalmopathy; osteoclastoma; bacterial infection; bone; ds.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 116..244
FT FT /*tag= a
FT FT /product= "secreted protein"
FT FT /transl_except= (pos:197..199, aa:Xaa)
FT FT /note= "Xaa=unknown"
FT sig_peptide 116..178
FT mat_peptide /*tag= b
FT FT 179..241
FT FT /*tag= c
PN MO9845712-A2.
PD 15-OCT-1998.
XX
PF 07-APR-1998; 98WO-US06801.
XX
PR 30-MAY-1997; 97US-0048184.
PR 08-APR-1997; 97US-0042726.
PR 08-APR-1997; 97US-0042727.
PR 08-APR-1997; 97US-0042728.
PR 08-APR-1997; 97US-0042754.
PR 08-APR-1997; 97US-004825.
PR 30-MAY-1997; 97US-0048068.
PR 30-MAY-1997; 97US-0048070.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Feng P, Ni J, Rosen CA, Ruben SM, Yu G;
XX WPI: 1998-594496/50.
XX P-PSDB: AAM83933.
XX
DR New isolated human genes and secreted polypeptide(s) they encode -
XX PT useful for the diagnosis and treatment of e.g. cancers, CNS
XX PT disorders, immune system disorders, inflammatory disease and
XX PT bacterial infections
XX
PS Claim 4: Pages 102-103; 142pp; English.
XX
XX This sequence represents a nucleic acid molecule designated Gene 3 from
XX the human cDNA clone HE2FV03 (deposited as clone ATCC 97955 and ATCC
XX 209074) which encodes a secreted human protein. This gene is expressed
XX primarily in infant brain, prostate, embryo and to a lesser extent in
XX parathyroid, adrenal gland, tumour and thymus and is useful for
XX diagnosing and treating CNS and immune-system diseases.
XX The invention relates to 20 novel genes and their fragments (AAV69611 to
XX AAV69630) and corresponding secreted proteins (AAW83931 to AAW83950)
XX which are useful for preventing, treating or ameliorating medical
XX conditions e.g. by protein of gene therapy. Also pathological conditions
XX can be diagnosed by determining the amount of the new polypeptides in a
XX sample or by determining the presence of mutations in the
XX polynucleotides. Specific uses are based on which tissues they are most
XX highly expressed in, and include developing products for the diagnosis or
XX treatment of central nervous system (CNS) and immune system diseases,
XX reproductive disorders, cancers, congenital malformations, degenerative
XX diseases, trauma, inflammatory disease, neoplasia, metabolic disorders,
XX diseases in testes, placenta, liver, brain and activated T cells, spleen
XX diseases, lung diseases, heart diseases, rhabdomyosarcoma and disorders
XX of the endocrine system or other endocrinopathies, e.g. endocrine
XX polyglandular syndrome, endocrinoma, and endocrine ophthalmopathy,
XX osteoclastoma and other bone remodelling disorders, bacterial infections
XX and sepsis. The polypeptides are also useful for identifying their
XX binding partners.
XX
XX Sequence 2067 BP; 639 A; 363 C; 407 G; 651 T; 7 other:

```


XX	AAZ29971:
AC	
XX	
DT	06-JUL-1999 (first entry)
XX	
DE	PDZ domain-containing protein gene FH950.
XX	
KW	PDZ domain; gene expression: human umbilical vascular endothelial cell;
KM	HUVEC; stimulation; tumour necrosis factor; TNF; protein binding; PCR;
KX	cell; proliferation disorder; cancer; primer; amplification; ss.
OS	Homo sapiens.
PN	M09907846-A1.
PD	18-FEB-1999.
XX	
PF	12-AUG-1998; 98WO-JP03603.
XX	
PR	19-JUN-1998; 98JP-0189944.
PR	12-AUG-1997; 97JP-0230356.
PA	(CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX	
PI	Funahashi S, Miyata S;
XX	
DR	WPI; 1999-167423/14.
DR	P-PSDB; AAY04740.
PT	Protein containing PDZ domain, whose expression is enhanced by TNF
PT	stimulation - plays an important role in protein/protein
PT	interactions and is used for screening for proteins for use in
PT	treatment of cell proliferation disorders such as cancer
XX	
PS	Example 7; Page 155-158; 240pp; Japanese.
XX	
CC	This sequence represents clone FH950 which encodes a new protein
CC	containing PDZ domains whose expression in human umbilical vascular
CC	endothelial cells (HUVEC) is enhanced by stimulation with tumour
CC	necrosis factor (TNF) alpha. The new protein is used to identify
CC	proteins which bind to it (particularly to the PDZ domains) and the
CC	genes encoding them, for use in the treatment of cell proliferation
CC	disorders such as cancer.
SO	Sequence 959 BP; 295 A; 198 C; 239 G; 227 T; 0 other;
Query Match	15.0%; Score 730.4; DB 20; Length 959;
Best Local Similarity	99.9%; Pred.No. 1.5e-194;
Matches 731; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
OY	1310 ttcctctgcttaccacgaattcagttcgaaagggtcttgagacctgtcttgacaagagt 1369
Dd	1 ttcccttcgtctcacccgacttcagctggaaaaggctctgagtaaccgtctgaacsagact 60
OY	1370 cccttgacctgaatgcgtgagtgctgcaagcttcaaatagtatctaagaatctttgaaa 1429
Dd	61 cccitggcctgtaaatgctgagtgctgcatgctccaataatgatctaagaatctttgaaa 120
OY	1430 ggacctataatatagcaaaaaggcaattcagccttagaagtgacagtgtagtctaataag 1489
Dd	121 ggaactattaataaacaaaaggcaattcagccttagaagtgacagtgtagtctaataag 180
OY	1490 atgagcttggagatgactgttcgaaagcatattcatcagagaggtgccattgtcagatggcc 1549
Dd	181 atgagcttggagatgactgttcgaaagcatattcatcagagaggtgccattgtcagatggcc 240
OY	1550 ggaatgcacattggggagacgcgcactgttgtccatlaagaaagagctccacacagtgtaacca 1609
Dd	241 ggaatgcacattggggagacgcactgttgtccatlaagaaagagctccacacagtgtaacca 300
OY	1610 atgcccaaggacagactatgtttgagaagaacattctcattatggcccctgacataaaactta 1669
Dd	

Db	301	atgccacgacgcgacatgttgaagaagaacattctctatcttgccctgacataaataa	360
QY	1670	cttatgtgctgcacgaacatttggaaaggtccaataaagcttggagacaacacttgaa	1729
Db	361	cttatgtgctgcacgaacatttggaaaggtccaataaagcttggagacaacacttgaa	420
QY	1730	gagcaatgacactgataattttcttcatacactgacagagacattccaattaccag	1789
Db	421	gagcaatgacactgataattttcttcatacactgacagagacattccaattaccag	480
QY	1780	agcagaagaagggaagggtgaagaagacgaactccaacaacagcatatagcaattga	1849
Db	481	agcagaagaagggaagggtgaagaagacgaactccaacaacagcatatagcaattga	540
QY	1850	atcagcccaacgcgggtggaactcttgagagaaacaaacgaactccttagcatcaacatg	1909
Db	541	atcagcccaacgcgggtggaactcttgagagaaacaaacgaactccttagcatcaacatg	600
QY	1910	ttgtgtgacgaagggaatgaggagtcggcctaagcaatgagagaatgtaggggcatctta	1969
Db	601	ttgtgtgacgaagggaatgaggagtcggcctaagcaatgagagaatgtaggggcatctta	660
QY	1970	tcaaacatgtcttggaagatagtcacgcgctggcgaataatgaaaccttgaaactggagta	2029
Db	661	tcaaacatgtcttggaagatagtcacgcgctggcgaataatgaaaccttgaaactggagta	720
QY	2030	gaatcgttagagg 2041	
Db	721	gaatcgttagagg 732	

RESULT 13
AAK29967
ID AAK29967 standard; DNA: 1776 BP.
AC AAK29967;
XX 06-JUL-1999 (first entry)
DE PDZ domain-containing protein gene clone D-2.
XX
XX PDZ domain; gene expression; human umbilical vascular endothelial cell;
KW HUVEC; stimulation; tumor necrosis factor; TNF; protein binding; PCR;
XX cell; proliferation disorder; cancer; primer; amplification; ss.
OS Homo sapiens.
XX
XX W09907846-A1.
XX
XX 18-FEB-1999.
XX
XX PD 12-AUG-1998; 98WO-JP03603.
XX
XX PR 19-JUN-1998; 98JP-0189944.
XX PR 12-AUG-1997; 97JP-0230356.
XX
XX PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
XX Funahashi S, Miyata S;
XX
XX WPI: 1999-167423/14.
XX DR P-PSDB: AAY04746.
XX
XX
XX PT Protein containing PDZ domain, whose expression is enhanced by TNF
XX PT stimulation - plays an important role in protein/protein
XX PT interactions and is used for screening for proteins for use in
XX PT treatment of cell proliferation disorders such as cancer
XX
XX Example 7; Page 140-146; 240pp; Japanese.
XX
XX This sequence represents clone D-2 which encodes a new protein containing
XX PDZ domains whose expression is enhanced by stimulation with tumor necrosis factor
XX cells (HUVEC) is enhanced by stimulation with tumor necrosis factor

(TMF) alpha. The new protein is used to identify proteins which bind to it (particularly to the PDZ domains) and the genes encoding them, for use in the treatment of cell proliferation disorders such as cancer.

Sequence 1776 BP; 558 A; 345 C; 449 G; 424 T; 0 other:

Query Match 11.5%; Score 563.6; DB 20; Length 1776;
Best Local Similarity 99.3%; Pred. No. 1.4e-147;
Matches 566; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

OY 212 aggttaatgataacttacttctggggaataaccacgaatgtgtgtaataatcttaaaag 271
    |||||||
DB 1207 aagtaaatgataacttacttctggggaataaccacgaatgtgtgtaataatcttaaaag 1266
OY 272 aactgcctataagaatgacaaatgtgtgtctgtcgtcgaactgtgcaccaccaccacat 331
    |||||||
DB 1267 aactgcctataagaatgacaaatgtgtgtctgtcgaactgtgcaccaccaccacat 1326
OY 332 cagaatttgatagctgctgactatgttatgtagctaaagaaagcctaacgtatgac 391
    |||||||
DB 1337 cagaatttgatagctgctgactatgttatgtagctaaagaaagcctaacgtatgac 1386
OY 392 taagtgaattcatcaggtcatcagaagacagatcagctgtgctgagatgactgacg 451
    |||||||
DB 1387 taagtgaattcatcaggtcatcagaagacagatcagctgtgctgagatgactgacg 1446
OY 452 gtccagaatcagaagaaggttcaagacacttggccatgtggaggtcgtgcatcagaca 511
    |||||||
DB 1447 gtccagaatcagaagaaggttcaagacacttggccatgtggaggtcgtgcatcagaca 1506
OY 512 tagagctgagaagaaggggcaaaagactgtgttagacatttaagattatacagaatccaa 571
    |||||||
DB 1507 taatgctgagaagaaggggcaaaagactgtgttagacatttaagattatacagaatccaa 1566
OY 572 ttgaccgcagcaagcactgtgataataatcgtcttctgtgctgcgcgcatctgtgaaa 631
    |||||||
DB 1567 ttgaccgcagcaagcactgtgataataatcgtcttctgtgctgcgcgcatctgtgaaa 1626
OY 632 aggatggaagacacttctccctgtgacgcacccaatgtttgtaaaagatgtaacttgaaa 691
    |||||||
DB 1627 aggatggaagacacttctccctgtgacgcacccaatgtttgtaaaagatgtaacttgaaa 1686
OY 692 acaggaagcttgagaagaactgtagaagacactgaaaggagccgcgaaggtctggaaga 751
    |||||||
DB 1687 acaggaagcttgagaagaactgtagaagacactgaaaggagccgcgaaggtctggaaga 1746
OY 752 taagagctgtaagccttaacccttcaac 781
    |||||||
DB 1747 taagagctgtaagccttaacccttcaac 1776

```

RESULT 14

AA86366: standard; DNA: 5836 BP.

AA86366:

29-SEP-1999 (first entry)

DNA encoding the human MMS1 protein.

Human: MMS1 protein; MMS1 interacting protein; tumour suppression;

KW MMS1 pathway; Immunogen; cancer; cell neoplastic growth; ss.

OS Homo sapiens.

Location/Qualifiers

115..5757

/tag- a

/product- MMS1_protein

/note- "no termination codon"

```

PN M09936566-A1.
XX 22-JUL-1999.
PD 19-JAN-1999; 99WO-US00995.
PF 20-JAN-1998; 98US-0071861.
PR (MYRI-) MYRIAD GENETICS INC.
PA Bartel PL, Tavtigian SV;
PI WPI: 1999-458472/38.
XX P-PSDB: AAY24025.
DR MMS1, an MMS1 (tumour suppressor) interacting protein and related
PT polynucleotides
PS Claim 2; Page 80-88; 107pp; English.

```

The present sequence encodes a MMS1 protein. The protein is a MMS1 interacting protein which is involved in tumour suppression activity in the MMS1 pathway. MMS1, antigenic fragments or fusion proteins of these are used as immunogens for antibody production. Primers derived from MMS1 genomic clones can be used for identification of MMS1 genes and for synthesis by amplification of MMS1 DNA or RNA. Detecting an alteration in MMS1 can be used to diagnose cancer. A germline alteration in an MMS1 gene is indicative of a predisposition to cancer. A somatic mutation in an MMS1 gene is indicative that the tissue is cancerous. Analysis of MMS1 and MMS1 (or PDZ domain 6 of MMS1) binding interactions can be used for detection of alterations in MMS1 associated with cancer. Wild-type MMS1 or a homologue can be used to supply wild-type MMS1 gene function (or a substantially similar function) to a cell, which has lost the gene function due to a MMS1 gene mutation. The gene suppresses neoplastic growth of the cell. Transgenic animals having an altered MMS1 can be used as a model for identifying drug candidates useful in treating cancer.

Sequence 5836 BP; 1754 A; 1215 C; 1444 G; 1423 T; 0 other:

Query Match 11.1%; Score 540.2; DB 20; Length 5836;
Best Local Similarity 63.4%; Pred. No. 1.1e-140;
Matches 887; Conservative 0; Mismatches 488; Indels 24; Gaps 3;

```

OY 3046 gatcctgaacactgcccattatccctgctgctgcaaaacacatcgagattccaaagg 3105
    |||||||
DB 4384 gaccgcgaacgctgcccattatccctgctgctgcaaaacacatcgagattccaaagg 4443
OY 3106 cgaagaagcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctt 3165
    |||||||
DB 4444 cgttaaggtctgctgctgctgctgctgctgctgctgctgctgctgctgctt 4503
OY 3166 atccatgaatttaagaagaagcagacatgttaagaatgaaagcttgagctgagat 3225
    |||||||
DB 4504 atccatgaatttaagaagaagcagacatgttaagaatgaaagcttgagctgagat 4563
OY 3226 cagatcttaagaagtgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 3285
    |||||||
DB 4564 cagatcttaagaagtgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaat 4623
OY 3286 gtccctgagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3345
    |||||||
DB 4624 gtccctgagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4683
OY 3346 aaagagagagagtgctgtaacacccctacatctgagctcagaaagacggaagagc 3405
    |||||||
DB 4684 cggagatgagagagtgagagtgagagtgagagtgagagtgagagtgagagtgagag 4743
OY 3406 ctgagatgaagtattgtgtgtaaaagaaagacatctgagatgtgtcagagatgtc 3465
    |||||||
DB 4744 ctgagcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4803

```

```

OY 3466 aaaggaagaaatgacagatcccgatgaaagactgtatccagggaagaccagataattgttg 3525
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4804 aaagcggaagcgagacgactgagatgagatgagatgagatgagatgagatgagatgagatg 4863
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3526 aatgggaagaaagctgtcgaatgagatgagatgagatgagatgagatgagatgagatg 3585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4864 aatgggaagaaagctgtcgaatgagatgagatgagatgagatgagatgagatgagatg 4923
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3586 ctgaagcagaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatg 3645
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4924 cagggaactgtgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacg 4983
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3646 aggcacatcacaacacacacacacacacacacacacacacacacacacacacacacac 3700
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4984 acgacacacacacacacacacacacacacacacacacacacacacacacacacacacac 5043
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3701 cactctctgatacgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcgaatcga 3760
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5044 ttcgctcctgacacacacacacacacacacacacacacacacacacacacacacacac 5103
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3761 catctgaata-----caggatlaagaaacagtcgaatgaaagagccct 3807
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5104 tcccgagaaatgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 5163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3808 actgaactcactggaatcagacacacacacacacacacacacacacacacacacacacac 3867
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5164 agtgaatcctcctggaatcagacacacacacacacacacacacacacacacacacacacac 5223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3868 atattatgacacacacacacacacacacacacacacacacacacacacacacacacac 3927
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5224 gtaattatgacacacacacacacacacacacacacacacacacacacacacacacacac 5283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3928 ggggaatgacacacacacacacacacacacacacacacacacacacacacacacacac 3987
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5284 ggaatcctcctggaatcagacacacacacacacacacacacacacacacacacacacac 5343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3988 gttacacacacacacacacacacacacacacacacacacacacacacacacacacac 4047
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5344 gttacacacacacacacacacacacacacacacacacacacacacacacacacacac 5403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 4048 gttgacacacacacacacacacacacacacacacacacacacacacacacacacacac 4107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5404 ataaagcgcacacacacacacacacacacacacacacacacacacacacacacacacac 5457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 4108 ctgaacacacacacacacacacacacacacacacacacacacacacacacacacacac 4167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5458 tcgcccacacacacacacacacacacacacacacacacacacacacacacacacacac 5517
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 4168 ctgaagcgcacacacacacacacacacacacacacacacacacacacacacacacacac 4227
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5518 ttggaagaaagcctcctggaagcctcctggaagcctcctggaagcctcctggaagcctc 5577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 4228 ggaacacacacacacacacacacacacacacacacacacacacacacacacacacacac 4287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5578 ggaacacacacacacacacacacacacacacacacacacacacacacacacacacacac 5637
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 4288 cgtctgaagaaagcgcacacacacacacacacacacacacacacacacacacacacacac 4347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5638 cgtatgaagaaagcgcacacacacacacacacacacacacacacacacacacacacacac 5697
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 4348 catgaagaaagcgcacacacacacacacacacacacacacacacacacacacacacacac 4407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5698 catgaagaaagcgcacacacacacacacacacacacacacacacacacacacacacacac 5757
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 4408 tcttgaattgctgcacaga 4426
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5758 tcatgagcctcgagccca 5776
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 15
 AAC27886
 ID AAC27886 standard; cDNA; 229 BP.
 XX

```

AC AAC27886;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 31961.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KM gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EPI033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 9905-0122487.
XX
PA (GENSET) GENSET.
XX
PI Dumas Mline Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 31961; 71bp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or poly(A) RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 229 BP; 99 A; 41 C; 27 G; 58 T; 4 other;
XX
Query Match 2.58; Score 124; DB 21; Length 229;
Best Local Similarity 96.9%; Pred. No. 9e-25;
Matches 124; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 4752 gctaaccaagaatgctcctgtaagtagtagtagtagtagtagtagtagtagtagtagtagtag 4811
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 gctaaccaagaatgctcctgtaagtagtagtagtagtagtagtagtagtagtagtagtagtag 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 4812 agaaatcctcctggaagcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 4871
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 agaaatcctcctggaagcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 4872 ccccaaac 4879
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 121 mcccccaam 128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: July 12, 2001, 21:53:11
 Job time: 10454 sec

Fri Jul 13 15:00:16 2001

us-09-502-698-3.rng

W

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2001, 18:56:47 ; Search time 112.56 Seconds
(without alignments)
8031.898 Million cell updates/sec

Title: US-09-502-698-3

Perfect score: 4880
Sequence: 1 cccgggccccggcgacagtg.....tctctcccccacccaact 4880

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 317530 seqs, 92630169 residues

Total number of hits satisfying chosen parameters: 635060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCITUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	1.4	8043	5	PCT-US94-09943-1
2	66	1.4	8119	4	US-09-290-640-45
3	65.6	1.3	7218	1	US-08-232-463-14
4	64.4	1.3	8040	1	US-08-596-291-1
5	64.4	1.3	8040	3	US-09-100-804-1
6	63.6	1.3	712	1	US-08-410-804-4
7	63.6	1.3	712	1	US-08-259-514-4
8	63.6	1.3	712	2	US-08-858-311-4
9	63.6	1.3	1830	1	US-08-410-804-2
10	63.6	1.3	1830	1	US-08-259-514-2
11	63.6	1.3	1830	2	US-08-858-311-2
12	56	1.1	1396	2	US-09-151-631-2
13	52.8	1.1	1042	4	US-09-276-531-77
14	49.4	1.0	2150	1	US-08-580-680-2
15	49.4	1.0	2150	1	US-08-480-156A-2
16	49.4	1.0	2150	1	US-08-354-961-2
17	49.4	1.0	2150	2	US-08-581-094-2
18	49.4	1.0	2150	2	US-08-580-665-2
19	49.4	1.0	2150	2	US-08-581-142-2
20	49.4	1.0	2150	2	US-08-581-103-2
21	49.4	1.0	2150	5	PCT-US94-05442A-2
22	49.4	1.0	2151	2	US-08-944-449-3
23	47.8	1.0	1178	4	US-09-091-405-1
24	46.6	1.0	2058	4	US-08-749-391-1
25	46.6	1.0	2058	4	US-09-390-200-1
26	42.6	0.9	393	2	US-08-944-449-1
27	42.4	0.9	5852	1	US-07-867-106-2

28	42.2	0.9	2389	1	US-08-123-161A-13	Sequence 13, Appl
29	42.2	0.9	2389	1	US-08-483-278-13	Sequence 13, Appl
30	42	0.9	2447	2	US-09-014-969-14	Sequence 14, Appl
31	40.4	0.8	7218	1	US-08-232-463-14	Sequence 14, Appl
32	40.2	0.8	4527	3	US-09-045-632-1	Sequence 1, Appl
33	39.8	0.8	19124	2	US-08-487-826B-13	Sequence 13, Appl
34	39.4	0.8	1474	4	US-08-821-994-64	Sequence 64, Appl
35	38.6	0.8	2040	3	US-08-997-445D-1	Sequence 1, Appl
36	38.4	0.8	51952	4	US-08-947-823-1	Sequence 1, Appl
37	37.6	0.8	642	3	US-08-946-026-33	Sequence 33, Appl
38	37.6	0.8	2026	2	US-08-993-228-3	Sequence 7, Appl
39	37.4	0.8	3705	5	PCT-US96-03940-7	Sequence 1, Appl
40	37.4	0.8	5648	5	PCT-US96-03940-1	Sequence 1, Appl
41	37.2	0.8	602	1	US-08-764-100-8	Sequence 8, Appl
42	37.2	0.8	642	1	US-08-764-100-13	Sequence 13, Appl
43	37.2	0.8	643	1	US-08-764-100-7	Sequence 7, Appl
44	37.2	0.8	2993	1	US-08-764-100-2	Sequence 2, Appl
45	37.2	0.8	2993	1	US-08-764-100-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
PCT-US94-09943-1
Sequence 1, Application PC/TUS9409943
GENERAL INFORMATION:
APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL
TITLE OF INVENTION: EXPRESSION OF NUCLEOTIDE SEQUENCES FOR NOVEL PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
STREET: 600 ATLANTIC AVENUE
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09943
FILING DATE: 01-SEP-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/115,573
FILING DATE: 01-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: TWOMEY, MICHAEL J.
REGISTRATION NUMBER: P-38,349
REFERENCE/DOCKET NUMBER: L0461/7000W0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/720-3500
TELEFAX: 617/720-2441
TELEX: 92-1742 EZEKEL
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8043 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: HOMO SAPIENS
FEATURE:

NAME/KEY: CDS
LOCATION: 78..7478
PCT-US94-09943-1

Query Match 1.4% Score 66; DB 5; Length 8043;
Best Local Similarity 58.8%; Pred. No. 7.5e-09;
Matches 114; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 4187 gcttagcttcagatagttggaagatagtcagccctcagagacttaccattatg 4266
DB 4150 GCTTGGGGAATGATGACCGGAGGTGATGACAGTGTGACATGTTGATATG 4209
QY 4247 ttaaaacagctgttcgaaaggagcagccctcgaagcagcgtctgaaaggcgatc 4306
DB 4210 TGAAGCTGTATTCCTCCAGGAGCAGCAGAGTGTGATGTGATGATTCACAAAGTATC 4269
QY 4307 agatcatgtctgtaacatgagcagctcagaagagtcacatgaagaagctgtgcca 4366
DB 4370 GCGTCTAGCTGTCAAGTGTAGTGTAGAGAGAGCCACCATAGCAAGCTGTGAAA 4329
QY 4367 tccttaacggaca 4380
DB 4330 CACTGAGAAATACA 4343

RESULT 2

US-09-290-640-45
Sequence 45; Application US/09290640
Patent No. 6204055
GENERAL INFORMATION:
APPLICANT: Dean, Nicholas M.
APPLICANT: Marcuseon, Eric G.
TITLE OF INVENTION: Antisense Compound Modulation of Fas Mediated Signaling
FILE REFERENCE: ISPH-0351
CURRENT APPLICATION NUMBER: US/09/290,640
CURRENT FILING DATE: 1999-04-12
NUMBER OF SEQ ID NOS: 85
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 45
LENGTH: 8119
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (64)..(7521)
PUBLICATION INFORMATION:
JOURNAL: FEBS Lett.
VOLUME: 337
ISSUE: 2
PAGES: 200-206
DATE: 1994-01-10
DATABASE ACCESSION NUMBER: D21209/Genbank
DATABASE ENTRY DATE: 1999-02-05
US-09-290-640-45

Query Match 1.4% Score 66; DB 4; Length 8119;
Best Local Similarity 58.8%; Pred. No. 7.5e-09;

Matches 114; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 4187 gcttagcttcagatagttggaagatagtcagccctcagagacttaccattatg 4266
DB 4193 gcttggggaagagtcagcagagtggtgaagagtgtagcagcagcgtgtgcatcattg 4252
QY 4247 ttaaaacagctgttcgaaaggagcagccctcgaagcagcgtctgaaaggcgatc 4306
DB 4253 tgaagagctgtatcccccagggagcagagctcgtatgtagaattccaaaaggtgctc 4312
QY 4307 agatcatgtctgtaacatgagcagctcagaagagtcacatgaagaagctgtgcca 4366
DB 4313 gcgtccacagctgtcaatgtgagtgtagttagaagagcaccacataagaagcgtgtgaaa 4372

QY 4367 tccttaacggaca 4380
DB 4373 cactgagaatata 4386

RESULT 3

US-08-232-463-14/c
Sequence 14; Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)863-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fis
US-08-232-463-14

Query Match 1.3% Score 65.6; DB 1; Length 7218;
Best Local Similarity 5.5%; Pred. No. 9.2e-09;
Matches 23; Conservative 232; Mismatches 161; Indels 0; Gaps 0;

QY 1865 aacattggaaggttcaaaataagcttgggacacaaatctggaaggtatggcactg 1744
DB 1456 AAGGATGAGAAGATTGTTGTAACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1397
QY 1745 atattttcttcatacagcagcagacattccagaattaccagagcgaagaaggag 1804
DB 1396 RRR 1337
QY 1805 aggttggaagaacgaaattcaaaacacagacataagcaattggaatcagcccgcg 1864
DB 1336 RRR 1277
QY 1865 tgaactctggagagaacacagcaatccttaggcatacagcatgtgtgtgagcagga 1924

1.38; Score 64.4; DB 1; Length 8040;

HYPOTHETICAL: NO

;; TITLE OF INVENTION: LYMPHOCYTE CHEMOATTRACTANT FACTOR AND
;; TITLE OF INVENTION: USES THEREOF
;; NUMBER OF SEQUENCES: 2
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Scully, Scott, Murphy & Presser
;; STREET: 400 Garden City Plaza
;; CITY: Garden City
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 11530-0299
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/480,156A
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Frank S. DIGI10
;; REGISTRATION NUMBER: 31,346
;; REFERENCE/DOCKET NUMBER: 98742Y
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (516) 742-4343
;; TELEFAX: (516) 742-4366
;; TELEX: 230 901 SANS UR
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2150 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; US-08-480-156A-2

Query Match 1.0%; Score 49.4; DB 1; Length 2150;
Best Local Similarity 61.1%; Pred. No. 0.00026;
Matches 80; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
QY 4240 attatgttaaacagctgttgcgaagggagcagctctgtgaagcggagctctgaaaagg 4239
DB 592 ATTACGGTTACAGAGGTCTTCCAAATGGCTGGCTCCAGGAAGGACTATTTCAGAG 651
QY 4300 ggcagtcagatcattgctcaatgagcagagctctagaaggagtcacccatgaagaagct 4359
DB 652 GGCATATGAGCTTCTTCATCAACGCAAGTCTCTCAAGGGAGCCACGACCATGATGCC 711
QY 4360 gtgcacatcct 4370
DB 712 TTGGCCATCT 722

Search completed: July 12, 2001, 21:46:52
Job time: 10205 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2001, 17:37:37 ; Search time 3448.6 Seconds
(without alignments)
13376.421 Million cell updates/sec

Title: US-09-502-698-3

Perfect score: 4880
Sequence: 1 cccgggccccggcgacagtg.....tctctcccccacccaact 4880

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1022815 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

EST: *
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

117: gb_est48:*
 118: gb_est49:*
 119: gb_est50:*
 120: gb_est51:*
 121: gb_est52:*
 122: gb_est53:*
 123: gb_est54:*
 124: gb_est55:*
 125: gb_est56:*
 126: gb_est57:*
 127: gb_est58:*
 128: gb_est59:*
 129: gb_est60:*
 130: gb_est61:*
 131: gb_est62:*
 132: gb_est63:*
 133: gb_est64:*
 134: gb_est65:*
 135: gb_est66:*
 136: gb_est67:*
 137: gb_est68:*
 138: gb_est69:*
 139: gb_est70:*
 140: gb_est71:*
 141: gb_est72:*
 142: gb_est73:*
 143: gb_est74:*
 144: gb_est75:*
 145: gb_est76:*
 146: gb_est77:*
 147: gb_est78:*
 148: gb_est79:*
 149: gb_est80:*
 150: gb_est81:*
 151: gb_est82:*
 152: gb_est83:*
 153: gb_est84:*
 154: gb_est85:*
 155: gb_est86:*
 156: gb_est87:*
 157: gb_est88:*
 158: gb_est89:*
 159: gb_est90:*
 160: gb_est91:*
 161: gb_est92:*
 162: gb_est93:*
 163: gb_est94:*
 164: gb_est95:*
 165: gb_est96:*
 166: gb_est97:*
 167: gb_est98:*
 168: gb_est99:*
 169: gb_est100:*
 170: gb_est101:*
 171: gb_est102:*
 172: gb_est103:*
 173: gb_est104:*
 174: gb_est105:*
 175: gb_est106:*
 176: gb_est107:*
 177: gb_est108:*
 178: gb_est109:*
 179: gb_est110:*
 180: gb_est111:*
 181: gb_est112:*
 182: gb_est113:*
 183: gb_est114:*
 184: gb_est115:*
 185: gb_est116:*
 186: gb_est117:*
 187: gb_est118:*
 188: gb_est119:*
 189: gb_est120:*

190: gb_est110:*
 191: gb_est111:*
 192: gb_hic:*
 193: em_gss_fun:*
 194: em_gss_hum1:*
 195: em_gss_hum2:*
 196: em_gss_hum3:*
 197: em_gss_hum4:*
 198: em_gss_hum5:*
 199: em_gss_hum6:*
 200: em_gss_hum7:*
 201: em_gss_hum8:*
 202: em_gss_hum9:*
 203: em_gss_inv1:*
 204: em_gss_inv2:*
 205: em_gss_inv3:*
 206: em_gss_inv4:*
 207: em_gss_inv5:*
 208: em_gss_inv6:*
 209: em_gss_inv7:*
 210: em_gss_inv8:*
 211: em_gss_inv9:*
 212: em_gss_inv10:*
 213: em_gss_inv11:*
 214: em_gss_inv12:*
 215: em_gss_inv13:*
 216: em_gss_inv14:*
 217: em_gss_inv15:*
 218: em_gss_inv16:*
 219: em_gss_inv17:*
 220: em_gss_inv18:*
 221: em_gss_inv19:*
 222: em_gss_inv20:*
 223: em_gss_inv21:*
 224: em_gss_inv22:*
 225: em_gss_inv23:*
 226: em_gss_inv24:*
 227: em_gss_inv25:*
 228: em_gss_inv26:*
 229: em_gss_inv27:*
 230: em_gss_inv28:*
 231: em_gss_inv29:*
 232: em_gss_inv30:*
 233: em_gss_inv31:*
 234: em_gss_inv32:*
 235: em_gss_inv33:*
 236: em_gss_inv34:*
 237: em_gss_inv35:*
 238: em_gss_inv36:*
 239: em_gss_inv37:*
 240: em_gss_inv38:*
 241: em_gss_inv39:*
 242: em_gss_inv40:*
 243: em_gss_inv41:*
 244: em_gss_inv42:*
 245: em_gss_inv43:*
 246: em_gss_inv44:*
 247: em_gss_inv45:*
 248: em_gss_inv46:*
 249: em_gss_inv47:*
 250: em_gss_inv48:*
 251: em_gss_inv49:*
 252: em_gss_inv50:*
 253: em_gss_inv51:*
 254: em_gss_inv52:*
 255: em_gss_inv53:*
 256: em_gss_inv54:*
 257: em_gss_inv55:*
 258: em_gss_inv56:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1280.8	26.2	2756	192	AK019164	AK019164 Mus muscu
2	586.8	12.0	632	32	AV691714	AV691714 Mus muscu
3	581.2	11.9	634	32	AV695000	AV695000 AV695000
4	577.7	11.8	628	32	AV697648	AV697648 AV697648
5	573.8	11.8	616	32	AV688922	AV688922 AV688922
6	571.2	11.7	616	32	AV682916	AV682916 AV682916
7	570.2	11.7	615	32	AV696967	AV696967 AV696967
8	569.6	11.7	622	32	AV690061	AV690061 AV690061
9	569	11.7	582	103	AI879075	AI879075 AV690061
10	568.4	11.6	855	168	BE675982	BE675982 602083939
11	554.6	11.4	701	162	BE263698	BE263698 601192138
12	554.4	11.4	672	32	AV646633	AV646633 AV646633
13	546.4	11.2	670	32	AV646631	AV646631 AV646631
14	540	11.1	670	32	AV646666	AV646666 AV646666
15	531.2	10.9	667	32	AV646567	AV646567 AV646567
16	526.4	10.8	639	32	AV646667	AV646667 AV646667
17	516	10.6	631	32	AV646613	AV646613 AV646613
18	515.4	10.6	518	112	AM160850	AM160850 au76d11.Y
19	504	10.3	516	105	AL138114	AL138114 DKFZP547D
20	497	10.2	497	112	AM161163	AM161163 au79c01.Y
21	491.2	10.1	499	156	C75629	C75629 C75629 Huma
22	489	10.0	585	140	BE780556	BE780556 601468856
23	484	9.9	606	32	AV693692	AV693692 AV693692
24	455.4	9.3	459	15	AI053851	AI053851 q116c02.s
25	446.4	9.1	457	5	AA18329	AA18329 EST20397
26	443.4	9.1	607	122	AM958106	AM958106 ESR370176
27	443.4	9.1	609	122	AM958109	AM958109 ESR370179
28	437.6	9.0	621	32	AV646668	AV646668 AV646668
29	434	8.9	434	112	AM163661	AM163661 au85901.Y
30	431	8.8	789	137	BE573846	BE573846 601331826
31	430.8	8.8	599	188	T26957	T26957 17n20p0r0.C
32	426	8.7	448	112	AM160606	AM160606 au67f05.Y
33	419	8.6	440	24	AI750090	AI750090 at35f11.x
34	416.2	8.5	456	142	BE928273	BE928273 QV4-CT049
35	414	8.5	414	22	AI563447	AI563447 tq7a09.x
36	413.4	8.5	417	2	AA132116	AA132116 z006g02.s
37	411.2	8.4	433	103	AI922328	AI922328 w90h010.x
38	410.8	8.4	421	112	AM157481	AM157481 au79c01.x
39	405	8.3	438	2	AA132350	AA132350 z006g02.r
40	400.2	8.2	860	175	BE293915	BE293915 602391342
41	395.6	8.1	601	151	BF608121	BF608121 MY1-00113
42	381	7.8	395	7	AA450257	AA450257 z442g04.r
43	380.8	7.8	581	153	BG404830	BG404830 602420151
44	345.4	7.1	355	9	AA568420	AA568420 n116c04.s
45	340.6	7.0	443	151	BF652538	BF652538 276126 MA

ALIGNMENTS

RESULT	1																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
--------	---	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

```

/codon_start=2
/protein_id="BAB31579.1"
/db_xref="GI:12859228"
/translation="PGDHTLAVDEVVAGCPVEKFLSLTKATATVLTVAEPAP
AVPSSATVSGERKDNSTPAVPAPDLEPSPSRSSSTPAVPADPTCPTEPCEET
IEISKGTGLSLTVGSDTLGLIITHEYTEBAKDKGLMAGDDILEVNIIDLEK
ATMDEALNVRKQEPQVRLTLRDAPYKEDVDYDTTIELQKPPGGLISVGRN
DTGVFSDIYKGGIADADGRIMQDQILMNGEDVRATQEAVALKCSIGAVTLEY
GRVKAAPFHSERPSOSQVSESSLSFTPELPSINTSESLSKNALASEIQRRLR
TVEIKKPPASL"
BASE COUNT      820 a      550 c      659 g      727 t
ORIGIN

```

```

Query Match      26.2% Score 1280.8; DB 192; Length 2756;
Best Local Similarity 79.0%; Pred. No. 0;
Matches 1620; Conservative 0; Mismatches 377; Indels 53; Gaps 6;

OY 2776 ggagatcagatactgctgtagatgataatgltgtgttaccctatgaaaagttc 2835
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 5 GGAGATCACATCTTGCGTGTAGATGATGATGATGCTGGGTGCTTGGAAAGTTC 64

OY 2836 attagccttcgaagagacgcaagaatgacagtaaacctaccatcagctgagaatcca 2895
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 65 ATCAGCCTTCTGAAGACGGCAAGCACTGTAAACCTGACTGTCGAGCTGAGATCCA 124

OY 2896 gattccagagctgtctctcagcagctgtgacgacgttgagaaaaaagacagctcc 2955
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 125 GCTGTGTCCAGCTGTCTTCTTTCACGCTTAAACGTCAGTGGAGAAAGAACACACTCC 184

OY 2956 cagctctgtagtgcacagctcgtgtcccaagaaccggaagtcacatccgaatacaagc 3015
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 185 CAGACTCCTGCAAGTCCCA-----GCTCCAGACTGGAACCCATCCCAAGTCAAGC 235

OY 3016 agatcatcaacagcagcaatttgcctctgtagctgcaacctgaccatcccatatccctggc 3075
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 236 AGGTCCTCACACACGACAGTGTCTTGTCTTCAACCTGCTCCACTGCCCATCATCCAGGC 295

OY 3076 tgcgaacaacacatcagatctccaaagggcgacaagggctgtggccttgagatcgtltgg 3135
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 296 TGTGAGACAAATGTGAATTTCCAAAGCCAAACAGGCGCTGGGACATGAGTATTTGTGG 355

OY 3136 ggttcagacagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3195
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 356 GGCCTCAGACACTGCTGGGTGCTATTTATTCATGAAAGTTTATGAGAGGAGGACACAGC 415

OY 3196 tctaaagctgaagagctctgggctgagatcagatcttagaggtgaatggaatgacttg 3255
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 416 TGTAAAGTGGAAAGCTATGGGCTGGAGACACAGATTTTAGAGTAAATGGGATTGACTTG 475

OY 3256 aggaagggcacacatgataagcaatcaatgctctgagacagacgcaagagatgctgc 3315
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 476 CGAAGGCTTACACATGATGAAAGCAATCAATGCTGAGGACAGACCCCTCAAAAGATACGG 535

OY 3316 ctgacacctcagagatgagagcccatcaaaagagagagagatgtgtgaacctcact 3375
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 536 CTGACGCTCTACCGAGAGAGGCCCATCAAAAGAGAGATGCTGTGAAATCACTTCAAC 595

OY 3376 attagctgcaagaagaacgggaaagagcttagaataatattgtgtgaaagaagaac 3435
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 596 ATCAGAGCTGCAAGAGAGGCCAGCAAGGCGCTTGCGTGAATGATGTGGCAAAAGAAAT 655

OY 3436 galactagatattgtgtcagacatgtcacaagaggaatctgacatcccgatgaaaga 3495
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 656 GACACTGAGATATTGTATCAGACATTGTCAAGGAGGACATTTGAGACGCGGATGGGAGA 715

OY 3496 ctgacatcagagagacagataatctatgtgtagagagagagcgttcgtaactcccaa 3555
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 716 CTGATGCAAGGAGCAACATTTTAAATGCTGAATGAGAAATGTCCTCATGACCAACGAG 775

OY 3556 gaagcaggttgcagcttgcctaaagtgttccttagacagagaaacttgaagtggagaaga 3615
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 776 GAGGAGTGGCTGCTGCTTAAAGTTCCTAGGTGCAATACCTGTGAGAGTGGAGA 835

```

```

OY 3616 atcaaaagctgtccatccatctcagagagagggccatctcaaacccagcaggtgagtaa 3675
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 836 GTCAAAAGCTGCCCATTTCCACTCAGAACGAGGCCCTTCAAGAGAGCCAGTGAAGTAG 895

OY 3676 ggcagcgtctctcttccatcttccactctctgatacagtaactgtagctagtaa 3735
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 896 AGCAGCCCTGTATCTTCTACCTCCCTCATTTCTGTGAATTAATATCATGAGTCAATGGA 955

OY 3736 agtagctcaagaagaatgcatctgcatctgaataacagagattaaagaagctgaaatg 3795
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 956 ACTAATCTGAAGAAATAATGCAATTACATCTGAATTTCAAGAGCTAAGACAGTCAATA 1015

OY 3796 aaaaagggccctactgacatctggaaatcagcatcgtctgaggagtagagcccatc 3855
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1016 AAAAAAGGCGCTGCGATCTGCTGCTGATCGATCGATCGATGAGAGGAGCGGCCGCTC 1075

OY 3856 ggtgtagtgcctataattatgtcaatgatacgaaccaactgtgagtcagacagaccaa 3915
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1076 GGAGATGTCCCTAATTTATTTGCGATGATGCAACCAATGCTGTGACCTCAAAACACA 1135

OY 3916 aaactcagagttggagatgagattgtccacatctgtgtgacatccactgaaggatgact 3975
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1136 AAACCTCAGAGTTGGGATGAGATTTGTCACCATCTGTGGACATCCACGATGGAGTACT 1195

OY 3976 cacacccaagcagttaaacctactgaaaatgcatctgtctcatatgaaatgcaagtggt 4035
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1196 CACACACAGGGGGTTAATTATGATAAATAATGCTCAGGCTCATTTGAATGACAGTGGTT 1255

OY 4036 gctggaagagacgtgagtggtgtgtcaaggtcaatcaagagagcgtcgaatctcagctc 4095
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1256 GCTGAGAGAGATGTGAGTGTGTGTGTCACGGGTATCGCAAGCAACTGTCCTTGTGTT 1315

OY 4096 tcttcaactgtgagctgaacacagatactctcagagtagattagagacccctcaatgt 4155
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1316 GCTTTCACATGGGCTGACATCAAGCATGATTTTCCGATGATTTAGGCCCTCCACAGCT 1375

OY 4156 aagctcatlaacatgagcagagagacagatgagcttaaggctcagatagttggaagatat 4215
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1376 AAGACCATTTAGCTTAGACCGGAGACACAGATGCTTAGGCTTACAGCATGTGAGCGGCTAT 1435

OY 4216 ggcagccctctgtagagatcccatatgataatgtaaaagagttgcaaaaggagagacc 4275
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1436 GGCAGCCTCATGGAGACTTACCAATTTATGTGAATAAGTGTGGCAAAAGGAGACACA 1495

OY 4276 tctgaagacgagcgtctgaaagggcgatcagatcattgtgtgtcaatgagagagctca 4335
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1496 GCAGAGATGGGCGCTTAATAAAGGGCGATCAGATCATGCTGTCATAGGGCAAAAGCTTA 1555

OY 4336 gaagagtcacccaatgaaagagcgtgtgtcatccttaaacgaaacaaagcagctcact 4395
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1556 GAAGGAGTGACCCATGAAAGAGCTGTTCCATCCCAAGAGAGCAAAAGGACACCGTCAAC 1615

OY 4396 ttgtagtctctctcttgatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4455
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1616 CTCAATGTTCTCTGTGAAGTACTGCGAGCTGAACACCAACGACCGCTGCTCCCT 1675

OY 4456 cctactgtaagaagaatgcactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4515
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1676 CCTACTGTAAAGAGAGAGACTG-----TTGTATGCTGTGTGTGTGTGTGTGTGT 1725

OY 4516 caaaactgtaaggagaaataacacacttaagtctcttcttccatctagaatgtcttctc 4575
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1726 CTACAGTGGAGAGCGGAGAA-----AACACAGTGTGTGTGT 1758

OY 4576 tactgacaacctaatcaatcttcttcttcttcttcttcttcttcttcttcttcttctt 4630
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1759 TATCTACAGCCCAACCATTTTCTTTACTCTGCAATTTATGATCATATATATCAATAA 1818

OY 4631 gaaggaatattgtgtgaatcgtctgttattgtgtgagatataatgtcttcttcttct 4690
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1819 GAAGAGATATTGCAATGATCAACCTCAGTTTATTTATCTGACA-ATATCTTAACAATTTAAG 1877

```

OY	4691	gtacatctggcgaagaatttatctcaagtcgaagtgcgtttagtataaagaagaatcatcttaa	4750
Dd	1878	gtcacctgtgacaaa-aattttttatatgtttcatcttgttactgtcgaaacaaaaatgatacca	1936
OY	4751	agctaccacaagaagaatggctccagtaagttagatlgaaanaatlgaanaatalataaataaag	4810
Dd	1937	agtttaggccatttagctttaagattaagatgcgaatttttgaaaaaaaagaagacgttttgacgtttta	1996
OY	4811	aagaaaatct 4820	
Dd	1997	taggaccttct 2006	
<div style="text-align: center;">}</div>			
RESULT	2		
LOCUS	AV691714		
DEFINITION	AV691714 GKC Homo sapiens cDNA clone GKCCME11 5', mRNA sequence.		
ACCESSION	AV691714		
VERSION	AV691714.1 GI:10293577		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 632)		
AUTHORS	Wu,T., Qian,B., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H. Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,D., Fu,G., Ren,S., Zhong,M., Lu,G., Yang .Y., Gu,Y., Chen,Z., and Han,Z.		
TITLE	Homo sapiens cDNA GK- clones		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Zequang Han		

FEATURES	source	location/Qualifiers
	1. .632	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="GKCM611"	
	/clone_11b="GKC"	
	/tissue_type="hepatocellular carcinoma"	
	/dev_stage="Adult"	
	/lab_host="SOLR"	
	/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"	
BASE COUNT	191 a	118 c 179 g 139 t 5 others
ORIGIN		

Query Match	12.0%	Score 586.8	DB 32	Length 632
Best Local Similarity	97.4%	Pred. No. 1,1e+148		
Matches	605	Conservative	0	Mismatches 12; Indels 4; Gaps 1
Qy	3171	tgaagttctatgaagaagagcgcacatgtaagaatggaagctctggctgagatcatgat	3230	
Db	1	tgaattttatgaaagagcagcagcatgtrtmaaatggamaaacctctggcctggagatcatgagt	60	
Qy	3231	cttagaaggttgaaatgaaatgacttgtagaagaagccacacatgtaagaacataaigtctct	3290	
Db	61	ctttagagctgaaatggaaattgacttttgaaagagccacacatgtaagaacataaigtctct	120	
Qy	3291	gagacagacgcacacagagagtgcgcctcgaactctcacaagaatgagccccacacaaaga	3350	
Db	121	gagacagacgcgcacagagagtgccgctgcacactttagacagatgagagcccccacaaaga	180	
Qy	3351	ggaggaagtgtgtgacacccctactactattgagtgtgcagaagaagccgggaaagcgcttag	3410	
Db	181	ggaggaagtggtgacacccctactactattgagtgtgcagaagaagccgggaaagcgcttag	240	

OY	3411	atgaatctgtgttgtaaagaacatactggatgatttctgtcaagatctgtcaagg	3470
Db	241	ATTAACTATTCTTGTGT-----GAACAGTACTGGAGTATTGTGTCAACATCTGTAAAGG	296
OY	3471	aggaattgcagatcccgatltgaaagactgataccagggagacacagatactatctgttgaaatg	3530
Db	297	AGGAATTGCAGATCCGATCGATGGAAGACTGATCGAGGAGACAGATATTAAATGGTAATGG	356
OY	3531	ggaaagaagtttcgttaabgcctcccaagaagcgglttcgcgcttgcctaaagtgttcacctg	3590
Db	357	GGAAAGAGCTTCGTATCTATCCACCCACAGAGGGGTGTGCCCTTGTCTAAATGTGTTCCCTTAGG	416
OY	3591	cacagtaacctgtgaagtltgaaagaatcaaaagctgtgtccatccatctcaaggagaagaagcc	3650
Db	417	CACAGTAACTTTGGAAATTTGGAAGAAATCANAGCTTGTCATTTCCATTTCAGAAAGGAGGCC	476
OY	3651	atctcaaacaccagccaagttgagtgaaagcagcctgtctcttcttcacttcttccactctcgtg	3710
Db	477	ATTCANAGGACAGCCAGGTGAGTGAAGGACACCTGTCATCTMTCACATTTTCCACTCTCTGG	536
OY	3711	atccagtaacatcttgatgtcaacttggaagaagtacataaagaagaatgcattggcactcgaat	3770
Db	537	ATTCAGTACACTCTAGTCACCTGGAAATACCTCANAGGAAGATGCATTGGCATCTGGAAT	596
OY	3771	acaggatltgaagaacagtga	3791
Db	597	ACAGGATTAGAACAGTCA	617

FEATURES	RESULT	3
LOCUS	AV6955000	634 bp mRNA EST 25-SEP-2000
DEFINITION	AV6955000 GKC Homo sapiens cDNA clone GKCCMH01 5', mRNA sequence.	
ACCESSION	AV6955000	
VERSION	AV6955000.1	GI:10296863
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
AUTHORS	Wu,T., Qian,B., Huang,Q., Kang,B., Gao,X., Xu,Z., Xiao,X., Li,N., Peng,T., Liu,F., Qu,J., Song,H., Cheng,Z., Zeng,J., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.	
TITLE	Homo sapiens cDNA GK- clones	
JOURNAL	Unpublished (2000)	
COMMENT	Contact: Zequang Han Chinese National Human Genome Center at Shanghai 351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919 (ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn This clone is available at CHGC in Shanghai.	
	Location/Qualifiers	

FEATURES	source	Location/Qualifiers
	1..634	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="GKCMH01"	
	/clone_id="GRC"	
	/tissue_type="hepatocellular carcinoma"	
	/dev_stage="Adult"	
	/lab_host="SOLR"	
	/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"	
BASE COUNT	193 a	120 c
ORIGIN	179 g	140 t
		2 others
Query Match	11.98;	Score 581.2; DB 32; Length 634;
Best Local Similarity	97.78;	Pred No. 3.8e-147;

Matches 600; Conservative 0; Mismatches 10; Indels 4; Gaps 1;

OY 3171 tgaatttatgaagaagagagacgacatgtaagaatggaagactctggctgagatcagat 3230
 |||||
 DB 1 TGAAGTTTATGAGAGAGAGACGACATGTAAGATGGAAGACTCTGGCTGGAGATCAGAT 60
 OY 3231 cttaagaggatgaatggaatgactctgaggaagccacacatgataagaatcattctc 3290
 |||||
 DB 61 CTTAAGAGTGAATGGAATGTAAGTGAAGAGCCACACATGATGAAGCAATCATGTCTCT 120
 OY 3291 gagacagagccgacagagagagtgccctgacactctacagagatgagccctacaaaga 3350
 |||||
 DB 121 GAGACAGAGCCGACAGAGAGTGGCTGACACTTACAGAGATGAGGCCCTTACAAAGA 180
 OY 3351 ggaaggaagtgctgacacccctcactctgagctgagagagagccgggaaaaagcctag 3410
 |||||
 DB 181 GGAGAGAGTGTGTGACACCTCTCATTTAGCTGAGAGAGAGCCGGGAAAAAGCCTTAG 240
 OY 3411 attaaagtctgtgtaaaagaagacatactgagatctgtgtcagaacattgtcaagg 3470
 |||||
 DB 241 ATTAAGTATGTTGCTGCT---GAAACGATACCTGAGATTTGTGTGACACATTGTCAAGG 296
 OY 3471 aggaattcagaatcccgatggaagactatccagagagagacagatatattgtgtaagg 3530
 |||||
 DB 297 AGGAATTCAGATGCGCATGGAAGACTGTGAGGAGAGCCAGATATTAATGTAATGG 356
 OY 3531 ggaagagctgctgaatgctccccaagaagcggctgctgctgctcaaggctccttag 3590
 |||||
 DB 357 GGAAGACGCTGCTGTAATGCGACCCAGAGAGCGGTTGCCCTTGTAAAGTGTCCCTTAG 416
 OY 3591 caagatgaactctggaagtctggaagaatcaaaagctgtgctcattcgaatgaagaagcc 3650
 |||||
 DB 417 CACAGTAACTCTTGGAAGTGGAAAGATCAAAAGCTGTCCATTCATTCAGAGAGAGGCC 476
 OY 3651 atctcaaacacagcaggtggaagcagcctgtctctcttccattccacctctcgg 3710
 |||||
 DB 477 ATCTCAAMGACACCGAGTGAGTGAAGGCAAGCGCTGTCACTTTCCACTCTCTGG 536
 OY 3711 atccagatcattctgactgactggaagatgactcaagaagaatgcatgcatcgaat 3770
 |||||
 DB 537 ATCCAGTACATCTGAGTCACTGGAAGTGAAGTCAANAGAAATGATGATGCGATCGAAT 596
 OY 3771 acagggattgaaga 3784
 |||||
 DB 597 ACAGGATTAACGA 610

RESULT 4

LOCUS AV697648 628 bp mRNA EST 25-SEP-2000
 DEFINITION AV697648 GKC Homo sapiens cDNA clone GKCA001 5', mRNA sequence.
 ACCESSION AV697648
 VERSION AV697648.1 GI:10299511
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 628)
 AUTHORS Wu,T., Qian,B., Huang,O., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Ren,S., Zhong,M., Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.

TITLE Homo sapiens CDNA GK- clones
 JOURNAL Unpublished (2000)
 COMMENT Contact: Zeyuang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919(ex.45)
 Fax: 86-21-50801922
 Email: hanzyg@chgc.sh.cn

This clone is available at CHGC in Shanghai.

FEATURES

Location/Qualifiers
 source 1..628
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="GKCA001"
 /clone_id="GKC"
 /tissue_type="hepatocellular carcinoma"
 /dev_stage="Adult"
 /lab_host="SOLR"
 /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 196 a 116 c 174 g 140 t 2 others
 ORIGIN

Query Match 11.8%; Score 577; DB 32; Length 628;
 Best Local Similarity 96.6%; Pred. No. 5.3e-146;
 Matches 600; Conservative 0; Mismatches 17; Indels 4; Gaps 1;

OY 3171 tgaatttatgaagaagagagacgacatgtaagaatggaagactctggctgagatcagat 3230
 |||||
 DB 1 TGAAGTTTATGAGAGAGAGACGACATGTAAGATGGAAGACTCTGGCTGGAGATCAGAT 60
 OY 3231 cttaagaggatgaatggaatgactctgaggaagccacacatgataagaatcattctc 3290
 |||||
 DB 61 CTTAAGAGTGAATGGAATGTAAGTGAAGAGCCACACATGATGAAGCAATCATGTCTCT 120
 OY 3291 gagacagagccgacagagagagtgccctgacactctacagagatgagccctacaaaga 3350
 |||||
 DB 121 GAGACAGAGCCGACAGAGAGTGGCTGACACTTACAGAGATGAGGCCCTTACAAAGA 180
 OY 3351 ggaaggaagtgctgacacccctcactctgagctgagagagagccgggaaaaagcctag 3410
 |||||
 DB 181 GGAGAGAGTGTGTGACACCTCTCATTTAGCTGAGAGAGAGCCGGGAAAAAGCCTTAG 240
 OY 3411 attaaagtctgtgtaaaagaagacatactgagatctgtgtcagaacattgtcaagg 3470
 |||||
 DB 241 ATTAAGTATGTTGCTGCT---GAAACGATACCTGAGATTTGTGTGACACATTGTCAAGG 296
 OY 3471 aggaattcagaatcccgatggaagactatccagagagagacagatatattgtgtaagg 3530
 |||||
 DB 297 AGGAATTCAGATGCGCATGGAAGACTGTGAGGAGAGCCAGATATTAATGTAATGG 356
 OY 3531 ggaagagctgctgaatgctccccaagaagcggctgctgctgctcaaggctccttag 3590
 |||||
 DB 357 GGAAGACGCTGCTGTAATGCGACCCAGAGAGCGGTTGCCCTTGTAAAGTGTCCCTTAG 416
 OY 3591 caagatgaactctggaagtctggaagaatcaaaagctgtgctcattcgaatgaagaagcc 3650
 |||||
 DB 417 CACAGTAACTCTTGGAAGTGGAAAGATCAAAAGCTGTCCATTCATTCAGAGAGAGGCC 476
 OY 3651 atctcaaacacagcaggtggaagcagcctgtctctcttccattccacctctcgg 3710
 |||||
 DB 477 ATCTCAAMGACACCGAGTGAGTGAAGGCAAGCGCTGTCACTTTCCACTCTCTGG 536
 OY 3711 atccagatcattctgactgactggaagatgactcaagaagaatgcatgcatcgaat 3770
 |||||
 DB 537 ATCCAGTACATCTGAGTCACTGGAAGTGAAGTCAANAGAAATGATGATGCGATCGAAT 596
 OY 3771 acagggattgaagaatcga 3791
 |||||
 DB 597 ACAAGATTTAGAACAGTGA 617

RESULT 5

LOCUS AV688922 616 bp mRNA EST 25-SEP-2000
 DEFINITION AV688922 GKC Homo sapiens cDNA clone GKCA002 5', mRNA sequence.
 ACCESSION AV688922
 VERSION AV688922.1 GI:10290785
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
AUTHORS
1 (bases 1 to 616)
Wu,T., Qian,B., Huang,Q., Kang,B., Gao,X., Xu,Z., Xiao,H.,
Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Zeng,L.,
Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang
Y., Gu,Y., Chen,Z. and Han,Z.
Homo sapiens CDNA GK- clones
Unpublished (2000)

JOURNAL
COMMENT
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzgehc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
Source
Location/Qualifiers
1..616
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GKCANCO2"
/clone_11b="GKC"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT
ORIGIN
191 a 117 c 174 g 134 t

Query Match 11.8%; Score 573.8; DB 32; Length 616;
Best Local Similarity 97.7%; Pred. No. 3.9e-145;
Matches 605; Conservative 0; Mismatches 7; Indels 7; Gaps 2;

3171 tgaatttataagaagagcagcatgtataagatggaagacctggcgctggagatcagat 3230
1 tgaatttataagaagagcagcatgtataagatggaagacctggcgctggagatcagat 60
3231 cttaagagtgtaagtgaattgacttggagaagccacacatagtatgaagcaatctct 3290
61 cttaagagtgtaagtgaattgacttggagaagccacacatagtatgaagcaatctct 120
3291 ggaagacagcgcacagagagtgctgctgacactctacagagatgggccccatacaaaa 3350
121 ggaagacagcgcacagagagtgctgctgacactctacagagatgggccccatacaaaa 180
3351 ggaagacagcgcacagagagtgctgctgacactctacagagatgggccccatacaaaa 3410
181 ggaagacagcgcacagagagtgctgctgacactctacagagatgggccccatacaaaa 240
3411 attaaagtattgttgtaaaagaacagatctggagatttggtaagatgtgcaagg 3470
241 attaaagtattgttgtaaaagaacagatctggagatttggtaagatgtgcaagg 296
3471 aggaattgagatcccgatgagaagctgctcagagagagccagatatatttggtaagtg 3530
297 aggaattgagatcccgatgagaagctgctcagagagagccagatatatttggtaagtg 356
3531 ggaagacagcgcacagagagtgctgctgacactctacagagatgggccccatacaaaa 3590
357 ggaagacagcgcacagagagtgctgctgacactctacagagatgggccccatacaaaa 416
3591 caagatccttggaagtggaagatcaaaagctgctcattccattcagagaagagcgc 3650
417 caagatccttggaagtggaagatcaaaagctgctcattccattcagagaagagcgc 476
3651 acttaaaacccagcaggtgagagagagcctgctcttccattcctccactctctg 3710
477 attctaaagaccacaggtgagagagagcctgctcttccattcctccactctctg 536
3711 atccagatcatctgagctcagagagagtgctcagagaagatgcatgcatctgaat 3770

537 ATCCACTACATCTGACTCTCTGGAAGTAGCTCAAGAAAGATGCAATGGCATCTG--T 593
3771 acaggatataagacagtc 3789
594 ACAGGATTAAGAACAGTC 612

RESULT 6
AV682916 616 bp mRNA EST 25-SEP-2000
LOCUS AV682916 GKC Homo sapiens CDNA clone GKCCNH07 5', mRNA sequence.
DEFINITION AV682916
ACCESSION AV682916 GI:10284779
VERSION AV682916.1 GI:10284779
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS
1 (bases 1 to 616)
Wu,T., Qian,B., Huang,Q., Kang,B., Gao,X., Xu,Z., Xiao,H.,
Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Zeng,L.,
Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang
Y., Gu,Y., Chen,Z. and Han,Z.
Homo sapiens CDNA GK- clones
Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzgehc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
Source
Location/Qualifiers
1..616
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GKCCNH07"
/clone_11b="GKC"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT
ORIGIN
192 a 116 c 175 g 133 t

Query Match 11.7%; Score 571.2; DB 32; Length 616;
Best Local Similarity 97.9%; Pred. No. 2e-144;
Matches 601; Conservative 0; Mismatches 8; Indels 5; Gaps 2;

3171 tgaatttataagaagagcagcatgtataagatggaagacctggcgctggagatcagat 3230
1 tgaatttataagaagagcagcatgtataagatggaagacctggcgctggagatcagat 60
3231 cttaagagtgtaagtgaattgacttggagaagccacacatagtatgaagcaatctct 3290
61 cttaagagtgtaagtgaattgacttggagaagccacacatagtatgaagcaatctct 120
3291 ggaagacagcgcacagagagtgctgctgacactctacagagatgggccccatacaaaa 3350
121 ggaagacagcgcacagagagtgctgctgacactctacagagatgggccccatacaaaa 180
3351 ggaagacagcgcacagagagtgctgctgacactctacagagatgggccccatacaaaa 3410
181 ggaagacagcgcacagagagtgctgctgacactctacagagatgggccccatacaaaa 240
3411 attaaagtattgttgtaaaagaacagatctggagatttggtaagatgtgcaagg 3470
241 attaaagtattgttgtaaaagaacagatctggagatttggtaagatgtgcaagg 296

3471 aggaattgcagatcccgatggaagactgataccagggaagaccagatatattgtgtaatg 3530
|||||
297 AGGAATTGCAGATCCCGATGGAAGACTGATGCAGGAGACAGATATTAAATGGTAATGG 3536
3531 ggaagaagtttgtaatgctcctcccaagaagcggttgcgcttgcctaaagtctccctagg 3530
|||||
357 GGAAGAGCTTCTGTAATGCTCCACCAAGAGCGGTGCGCTTGTCTAAAGTCTCCCTAGG 416
3591 cacagtaacctgtggaattggaagaaatcaaaagctgtccattccattccagaagaagagcc 3650
|||||
417 CACAGTAACCTTGGAGAGTGGAAAGATCAAAAGCTGTCATCTTCATTCCATGAGAGAGGCC 476
3651 atctcaaacaccagcaggtagtgaaggcaagcctgtctctcttccattccattccactctg 3710
|||||
477 ATCTCAAGCAGCAGCGAGTGAAGGAGCGCTGTCACTTTCATCTTCCACTCTCTGG 536
3711 atccagtaacatctgtaagctcctcccaagaagcggttgcgcttgcctaaagtctccctagg 3770
|||||
537 ATCCAGTACATCTGATGCTACCTGGAAGAGCTCAAGAGAAATGCA-TGGCATCTGAAT 595
3771 acaggagtaagaa 3784
|||||
596 ACAGGATTAAGCA 609

RESULT 7
AV696967 615 bp mRNA EST 25-SEP-2000
LOCUS AV696967 GKC Homo sapiens cDNA clone GKCNF08 5', mRNA sequence.
DEFINITION AV696967.1 GI:10298830
VERSION AV696967.1 GI:10298830
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 615)
AUTHORS Wu, T., Qian, B., Huang, Q., Huang, C., Kang, B., Gao, X., Xu, Z., Xiao, H.,
Xu, X., Li, N., Peng, Y., Liu, F., Qu, J., Song, H., Cheng, Z., Zeng, L.,
Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Yang
Y., Gu, Y., Chen, Z., and Han, Z.
Homo sapiens cDNA GK- clones
Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
source
1. 615
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GKCNF08"
/clone_id="GKC"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 191 a 115 c 174 g 134 t 1 others
ORIGIN
Query Match 11.7%; Score 570.2; DB 32; Length 615;
Best Local Similarity 97.7%; Pred. No. 3,7e-144;
Matches 600; Conservative 0; Mismatches 9; Indels 5; Gaps 2;
3711 tgaagttatgaagaagagagagatgtaagatggaagactcgggctgtagatcagat 3720
|||||
1 TGAAGTTATGAAGAAGAGAGAGATGTAAGATGGAAGACTGTGGCTGGAGATCAGAT 60

3331 cttagaagtgtaagtaattgacttgagaagagccacacatgatgaagcaatcaatgctc 3290
|||||
61 CTTAGAGTGAATGTAATGACTTGAGAAAGGCCACACATCATATACCATATGTCCT 120
3391 gagaacagcgcacacagagagtgccctgcaactctacagagatgagcccatcaaga 3350
|||||
121 GAGACACAGCCACAGAGAGTGCCTGACACTTACAGATGAGCCCATCAAGA 180
3351 ggaagaagtggtgtaacccctccactattgagctgcagaagaagccggaaagccctagg 3410
|||||
181 GGAGAAAGTGTGTACACCCCTCACTATTGAGCTCACAAGAGCCGGAAAGGCTTAGG 240
3411 attagaattgtgttaaaagacgatactggaatttgcacagacatgtcacaag 3470
|||||
241 ATTAAGTATTGTTGCT----GAAACGATACGAGATTTTGTGACACATGTCAAGG 296
3471 aggaattgcagatcccgatggaagactgataccagggaagaccagatatattgtgtaatg 3530
|||||
297 AGGAATTGCAGATCCCGATGGAAGACTGATGCAGGAGACAGATATTAAATGGTAATGG 356
3531 ggaagaagtttgtaatgctcctcccaagaagcggttgcgcttgcctaaagtctccctagg 3530
|||||
357 GGAAGAGCTTCTGTAATGCTCCACCAAGAGCGGTGCGCTTGTCTAAAGTCTCCCTAGG 416
3591 cacagtaacctgtggaattggaagaaatcaaaagctgtccattccattccagaagaagagcc 3650
|||||
417 CACAGTAA-CTTGGAGAGTGGAAAGATCAAAAGCTGTCATCTTCATTCCATGAGAGAGGCC 475
3651 atctcaaacaccagcaggtagtgaaggcaagcctgtctctcttccattccattccactctg 3710
|||||
476 ATCTCAANAGCAGCAGCGTGAAGGAGCGCTGTCACTTTCATCTTCCACTCTCTGG 535
3711 atccagtaacatctgtaagctcctcccaagaagcggttgcgcttgcctaaagtctccctagg 3770
|||||
536 ATCCAGTACATCTGATGCTACCTGGAAGAGCTCAAGAGAAATGCAATTGGCATGTAAT 595
3771 acaggagtaagaa 3784
|||||
596 ACAGGATTAAGCA 609

RESULT 8
AV690061 622 bp mRNA EST 25-SEP-2000
LOCUS AV690061 GKC Homo sapiens cDNA clone GKCAN09 5', mRNA sequence.
DEFINITION AV690061
VERSION AV690061.1 GI:10291924
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 622)
AUTHORS Wu, T., Qian, B., Huang, Q., Huang, C., Kang, B., Gao, X., Xu, Z., Xiao, H.,
Xu, X., Li, N., Peng, Y., Liu, F., Qu, J., Song, H., Cheng, Z., Zeng, L.,
Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Yang
Y., Gu, Y., Chen, Z., and Han, Z.
Homo sapiens cDNA GK- clones
Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
source
1. 622
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GKCAND09"

BASE COUNT	194 a	118 c	173 g	136 t	1 others
ORIGIN					
Query Match	11.7%;	Score 569.6;	DB 32;	Length 622;	
Best Local Similarity	97.0%;	Pred. No. 5.5e-144;			
Matches 592;	Conservative 0;	Mismatches 14;	Indels 4;	Gaps 1;	
QY 3171	tgaagttatgaagaagagcagcaatgtaagaatggaagactctggtctgagatcagat	3230			
Db 1	TGAAGTTATGTAAGAGAGAGCAGCATGTAAAGATGGAAGACCTCGGCGTGAGATCAGAT	60			
QY 3231	cttagagtgtaagtggaattgacttgtaggaagccacaaatggtgaagcaataatgctct	3290			
Db 61	CTTAGAGTGTAAAGTAATGTAAGTACTTGAGAAAGGCCACACACTGTATGAAGCATTAATGTCTT	120			
QY 3291	gagacagacgcacacagagagtgccgctcagacactctacagagatgagccccatacaaga	3350			
Db 121	GAGACAGAGCCGACAGAGAGTGCGGCTGACACTCTACAGATGAGATGAGCCCATACAAAGA	180			
QY 3351	ggaggaagtggtgtagacccctcaactattgagctgcagaagaacccgggaaagccctag	3410			
Db 181	GGAGGAGACTGTGTGACACCCCTCCTACTATTGAGCTGCAGAAAGCCGGGAAAGCCCTAG	240			
QY 3411	attaagtatctgtgttaaagaagaactagtagtatgtgtgtaagatatgtgtaag	3470			
Db 241	ATTAAAGTATGTGTGTAAAGAAAGATACGTGGAGTATTTGTTCAGACATTGTCAAAAG	300			
QY 3471	aggaattgcagatcccgatctggaagactgattccagagagagaccagatatattgtgtaatg	3530			
Db 301	AGGATTCGACAGTGCCTGCGATGGAAGACTGATGAGGAGAGACCGAGATATTAATGCTAATG	360			
QY 3531	ggaagaagtgctgtaatgctctcccaagaagcgggtgcgcgtctgtctaagtgctcctag	3590			
Db 361	GGAAAGCCTGTGTAATGCCACCCCAAGAGCGGTTGCCGTTGCTAAAGTGTTCCTAG	420			
QY 3591	cacagtaaccttggaagtggaagaatcaaaagctgtgcatctccattccattcagaagaagc	3650			
Db 421	CACAGTAACTCTTGGAAAGTGGAAAGATCAAAACCTGGTCCATTCCATTCAGAAAGGAGCC	480			
QY 3651	atctcaaacacccagcagtggaagtggaagcagcctgtcctcttccactttcgaactctg	3710			
Db 481	ATCTCAAAAGCAGCCAGGAGTAGAGTAAGGCGACGCTGTGCATCTTTCCACTCTCTGG	540			
QY 3711	atccagatcatctgtagtacttggaagaatagctcaagaagaatgtagatgtagatctgaaat	3770			
Db 541	ATCCAGTACATCTGAGTACTCTGAAAGTAGCTC-----GAGAAATGCATTGCATGGAAT	596			
QY 3771	acagggatca 3780				
Db 597	ACAGGATTA 606				
RESULT 9					
LOCUS	A1879075	562 bp	EST	23-AUG-1999	
DEFINITION	au44b02.y1 Schneider fetal brain 00004 Homo sapiens cDNA clone				
IMAGE:2518539	5' similar to TR:043798	043798	MULTI PDZ DOMAIN		
PROTEIN 1	;	MRNA	sequence.		
ACCESSION	A1879075				
VERSION	A1879075.1	GI:5553124			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
REFERENCE	1 (bases 1 to 582)				

FEATURES	SOURCE
1. 582	Location/Qualifiers
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/clone="IMAGE:2518539"	
/clone_lib="Schneider fetal brain 00004"	
/sex="male"	
/tissue_type="frontal lobe"	
/dev_stage="5 months post-conception"	
/lab_host="DH10B"	
/note="Organ: brain; Vector: pBluescript SK (Stratagene); Site:1: SalI; Site:2: XhoI; Double-stranded cDNA was prepared from human fetal brain tissue. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-GAGAGAGAGAGAGAGCTCAGAGATCTTAATTAATTAATCCCCCCCCC-3' and 3' adaptor sequence: 5'-GAGAGAGAGAGAGCTCAGAGATCTTAATTAATTAATCCCCCCCCC-3'. The library was size-selected for >0.5 kb inserts and has an average insert size estimated at 1.2 kb. This library was constructed using the CAP-trapper method for full-length enrichment and has not undergone amplification. Library was constructed by Dr. Claudio Schneider (LNCIB-Area Science Park, Trieste, Italy)."	
BASE COUNT	161 a 115 c 129 g 176 t 1 others
ORIGIN	
Query Match	11.7%; Score 569; DB 103; Length 582;
Best Local Similarity	99.7%; Pred. No. 7.8e-144;
Matches	580; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Db	1 ATATTTCAGGATGATTTAGGACCTCTCAATGATAGTATTACCTAGAGCGAGACCA 60
Qy	4123 atatttcagatgattgagacctctccaatgtaagtcctatatacactagagcgagacca 4182
Db	1 ATATTTCAGGATGATTTAGGACCTCTCAATGATAGTATTACCTAGAGCGAGACCA 60
Qy	4183 gatgcttaggcttcagatgattgtaagagataatgcaacccctcagaggaactaccat 4242
Db	61 GATGGCTTAGGCTTCAGATGATTGAGAGATATGGACCCCTCAAGGAGACTTACCACAT 120
Qy	4243 tatgttaaacacagtcgttcgcaaaaggagcagcctctgaaagacgagcgtctgaaaggggc 4302
Db	121 TATGTTAAACAGTCTTTCGCAAAAGGAGACAGCCTCTGAAAGAGGAGCGTCTGAAAAAGGGC 180
Qy	4303 gatcagatcatctgctgtaaatgagcagagctctagaaggagtcacccatgaaagacgtgt 4362
Db	181 GATCAGATCATCTGCTGTAATGGCGAGAGTCTAGAAAGAGTCAACCATGAACACTGT 240
Qy	4363 gccatcctctaaagcaaaaggacgtctcactttgtagtgcctctcttgaaattggcgcg 4422
Db	241 GCCATCCTCTTAAACGACAAAAGGCATCTGCACTTTGAATGATTCCTTGAATTTGGCTGC 300
Qy	4423 cagaattgacccaacccctctagagccactctcactgtctgtaaaagagatgcactgctgc 4482
Db	301 CAGAAATTGAACCAACCCCTAGCTCAGCTCTCTTACTGTTAAAGAGAAATGCACTGGTCC 360
Qy	4483 tgaacaatttatactgctgtctcagccgggtcttcaaaactgtaaggggggaataaacact 4542

Db	361	TCACATATTTTATGCTGTGTTCCAGCCGGGCTCTTCAAAACTGTAGGGGGGAAATTAACACTT	420
Qy	4543	aaattcttttcatactagaagaagcttcttctactgaacaacccaacgaacttttctt	4602
Db	421	AAAGTTCCTTTTCTATCTAGAAAGCTTCTTACTGTCACAACCTTACATCATTTTCTT	480
Qy	4603	tctctctgcatcttctgtaactaaagaagaatatttctgtagtlyt-aaatcgatt	4661
Db	481	TTCCTCTTGCAATTMTGTGAACCTTAAGAAGAAAGCAATATTGTGTAGGTGAATTCGGTTT	540
Qy	4662	ttatttgtagagatatctcaatgcttcttgatgtaactatgggcaa	4703
Db	541	TTATTTGTGAGAAATCTPAATGTTTGTGATGCATAGGCAA	582
RESULT	10		
LOCUS	BF675982		
DEFINITION	BF675982	855 bp	mRNA
ACCESSION	BF675982		EST
VERSION	BF675982.1		21-DEC-2000
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 855)		
JOURNAL	NIH-MGC http://mgc.ncl.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Stransberg, Ph.D.		
	Email: cgaabs-remail.nih.gov		
	Tissue Procurement: CLONETECH Laboratories, Inc.		
	CDNA Library Preparation: CLONETECH Laboratories, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: LICM1066 row: m column: 10		
	High quality sequence stop: 614.		
FEATURES			
source	1..855		
	location/Qualifiers		
	organism="Homo sapiens"		
	db_xref="taxon:9606"		
	clone_image:4248369"		
	clone_id="NIH-MGC_83"		
	lab_host="DH10B (TI phage-resistant)"		
	note="Organ: prostate. Vector: pDNR-LIB (Clontech);		
	site:1: SfiI (ggccgcctggcc); site:2: SfiI (ggccatattggc		

Accession	Source	Organism	Reference	Title	Journal	Comment
Db	61	AAGTA-TCCTTGTAAGAAACGATACGTGAGTA-TTGTCACACATTTGTCAAGGAGG	118			
Qy	3474	aattgcagatcccgatgtaagaacgcatccaggggagaccagatatatgtgtgaatggga	3533			
Db	119	AATTGCACATCCCATGGAGAACGTGATGCAGGGAGACCAGATATTAAATGATGGGGA	178			
Qy	3534	agaagcttcglaatgcctccccaagaagcgtgtgcgcctcttcgaagtgtcccttaggcac	3593			
Db	179	AGAGGTTGTGTAATGCCAACCCAGAAGGGGTTGCCGCTTTGGTAAAGTGTTCCCGAGGCAC	238			
Qy	3554	agtaaccttggaagtttggaaatccaagctgtgtccatctccatccagaagagggccatc	3653			
Db	239	AGTAACCTTGGAAATGGGAAATCAAAAGCTGGTCATTTCCATTTCACAGAGAGGCCATC	298			
Qy	3654	tcaaacccagcaggtagtgaaggcagcctgtctctcttcacctttccacctctgcatc	3713			
Db	289	TCAAGGACGCCAGGTGATGAAAGGACACCTGTATCTTTCACTTTTCCACTCTGTGATC	358			
Qy	3714	cagtaacatctgtagtcaetgtaagtagctcaaagaagaatgcatctggcatctgaataca	3773			
Db	359	CAGTACATCTGATGTCACCTGGAAAGTACTCTCAAAAGAAATGCAATGGGCATCTGAAATACA	418			
Qy	3774	ggagcttaagaaacagttcgaaatgtaaaaaggggccctcagctacatctggaatcagcatcgc	3833			
Db	419	GGGATTAGAACACAGTCGAAATGAAAAAGGCCCTTACTGACTCACTGGGAATCAGCATCGC	478			
Qy	3834	tgagaggatgaagcagccaccttggtagtgtgctctatacttaattgcaatgaatgcacccaac	3893			
Db	479	TGGAGAGATAGCGAGCCACCTGTTGTATGTGCCATATTTATTAATGAATGATGCAC--AA	536			
Qy	3894	tgaggttgcagcaagaagcccaaacctcagagttggggatagatgttccacctctgtg	3953			
Db	537	CTGAGTTGCAC--AAGACCCAAAAAATCAGATGGGATAGGATTTGCACCATTTGTGG	594			
Qy	3954	cacatcccatgaggagcatgcatcaccacccaagcagtttaacctactgtaaaaatgcatctg	4013			
Db	595	GACATTCACATCGAGAGGATGATCAACAACCCAGCAAGTTTAACCTACTGAAAAATG--GTTTG	652			
Qy	4014	ctccatgtgaatgcagatgtgtctctgtgaagaagacgtgtagtgtgtccagaggtcatca	4073			
Db	653	GTCATATGAAATGCAAGGAGGTGTGTGAGAGGAAGGTGA-TGTTGTCCCAAGGTATCAGGA	711			
Qy	4074	ggagcctgcaggtccacgctctcttccatcagcggcctgacg	4113			
Db	712	GAGAGCTGCAGATTCCAGTTTCTTAACTAGTGGGGGAGAGAG	751			
RESULT	11					
BE263698	701 bp	mRNA	EST	13-JUL-2000		
LOCUS	601192138F1	NIH_MGC_7	Homo sapiens	cdna	clone	IMAGE:3536024 5',
DEFINITION	mRNA sequence.					
ACCESSION	BE263698					
KEYWORDS	BE263698.1	GI:9137243				
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryotes: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Homiidae: Homo.					
AUTHORS	NIH-MGC	http://mgc.nci.nih.gov/				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)					
JOURNAL	Unpublished (1999)					
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cga@p5.femail.nih.gov Plate: LCM217 row: 1 column: 09 High quality sequence stop: 620. Location/Qualifiers 1. 701					
FEATURES	source					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
	/clone="IMAGE:3536024"					

/clone.lib="NH.MCC.7"
 /tissue.type="small cell carcinoma"
 /cell_line="MGC3"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Lung; Vector: pORF7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 215 a 144 c 194 g 148 t
 ORIGIN

Query Match 11.4%; Score 554.6; DB 165; Length 701;
 Best Local Similarity 99.1%; Pred. No. 7.1e-140;
 Matches 568; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
 QY 1665 aattctatgtgctgcgcgaacatttgaaagatccaataaagcttgggacacaacac 1724
 DB 67 AATTCTTATGTCCTGCGCAACATTGGAGAGATTCAAAATPAGCTTGGACACCAATC 126
 QY 1725 tggagagatgagcactgagatattttctcatcactgagagacattccagaatt 1784
 DB 127 TGGAGAGATGAGCGCATCGATATTTTCTTCATACACTGGCAGACATTCAGAAAT 186
 QY 1785 accagagcagagagagagagaggtgtaagaaagcgaactccaacacacacataagca 1844
 DB 187 ACCAGAGCGAGAGAGAGGAGCGTGAAGAAAGCAACTTCAAAACACAGCATATAGCAA 246
 QY 1845 ttggatcagcccgccgagctggaactctggagagacccaacaacattccttgatcag 1904
 DB 247 TTGGAATCAGCCCGAGCGGGGTGAACTCTGGAGAGAACCAACCAATCCTTAGGCATCAG 306
 QY 1905 catctctggtagcagagagatggagatcgctgagcgaatgagagagatgagagagat 1964
 DB 307 CATTTGGTGGACGAGGATGGGAGTGGCTTAACCAATGAGAAATATAGGGGCAAT 366
 QY 1965 ttctatcaaacatgctctggaagatagtcagctggcacaacatggaacttgaacctgg 2024
 DB 367 TTTCATCAAAACATGTTCTGGAAGATAGTCACACTGGCAAAATGGAACCTTGAACCTGG 426
 QY 2025 agatagatcgtgtagagagcccaagtcagtcagatcagagcagagagagagctccatgtg 2084
 DB 427 AATATGAATCGTAGAGGCCACCCAGTCAGATCAGAGTCAGAGCCAGAGAAAGGCTCCATTGTG 486
 QY 2085 cagctgccccccacccctccttcagccttcgcaaatgggtatgatacaacaacagtc 2144
 DB 487 CAGTGTGCCCCCACCCTCTCTTCACCTTTCGCCAAATGGGTAGATCACACACAGTC 546
 QY 2145 atctgcaagcaaaatctcacaagatgtgagacaagagagatgagttggttcaagctggaa 2204
 DB 547 ACTGCAAGCAAAATCTCTCAAGATGTGACAAAGAGATGAG-TTGGTTACAGCTGGA 605
 QY 2205 aaatatcagagagcggttatggaaccccaagcagg 2237
 DB 606 AATATCAGAGAGCGTTATGAACTTAACAGG 638
 RESULT 12
 AV646633 672 bp mRNA EST 07-SEP-2000
 LOCUS AV646633 GLOC Homo sapiens cDNA clone GLCAP10 3', mRNA sequence.
 DEFINITION AV646633
 ACCESSION AV646633
 VERSION AV646633.1 GI:9867647
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 672)

AUTHORS Qian,B., Wu,T., Huang,O., Kang,B., Gao,X., Xu,Z., Xiao,H.,
 Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J.,
 Zeng,L., Xu,S., Gu,M., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu
 ,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.
 TITLE Homo sapiens cDNA clone
 JOURNAL Unpublished (2000)
 COMMENT Contact: Zeguang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919(ex.45)
 Fax: 86-21-50801922
 Email: hanzg@hgc.sh.cn
 This clone is available at CHGC in Shanghai.

FEATURES
 Location/Qualifiers

Source

1. 672

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="GLCAP10"

/clone.lib="GLOC"

/tissue.type="Corresponding non cancerous liver tissue"

/dev_stage="Adult"

/lab_host="SOLR"

/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
 XhoI"

BASE COUNT 204 a 152 c 173 g 141 t 2 others
 ORIGIN

Query Match 11.4%; Score 554.4; DB 32; Length 672;
 Best Local Similarity 87.7%; Pred. No. 7.9e-140;
 Matches 663; Conservative 0; Mismatches 8; Indels 85; Gaps 2;

QY 2711 gtggagatcaataaagagcttaacagagcagtgagtagcagccagatgagcacta 2770
 DB 1 GTGGAGTCATCATTAAGACCTTAACAGACATGAGGCTACACCGAGATGACGACTCA 60
 QY 2771 aagtcggaatcaagatgagcgtgtagatgagaaatgttgttgtaacctatgaa 2830
 DB 61 AAGTCGAGATCAGATGAGCTGCTGATGATGAAATTTGTGTGATTCCTATTGAA 120
 QY 2831 agttatagcctcttgaaagacagcaaatgacagtaaacatccatcctgtgaga 2890
 DB 121 AGTTTATGCTTCTGAGACAGCAAAATGACATTAACATTCATTCATGCTGAGA 180
 QY 2891 atccagatcccaagcgtgtctcctcaagcagctgtagcagctgagagaaaaaagaa 2950
 DB 181 ATCCAGATTCACAGGCTGTCTTCAGCAGCTGTGACAGCCAGTGAAGAAAAAGAAC 240
 QY 2951 gctccagctctgtatggtcccaagctggtctcccaagaccgagatccatccgaata 3010
 DB 241 GCTCCAGCTCTGTGATGGTCCACAGCTGTGCTCCCAACCGAGTCCATTCGAAATA 300
 QY 3011 caagcagatcatcaaacacagcaatttgcctctgagatcctgcaacctgcccattatcc 3070
 DB 301 CAAGCAGATCATCAACACAGCAATTTTGTCTTGTATCTCTGCAACCTGCCCATTAATCC 360
 QY 3071 ctggctcgcaaaacacacatcgagatttccaaaggcgcaacagagctgagcctgagatcg 3130
 DB 361 CTGGCTCGCAAAACACCATCGAGATTTCAAAGGGCGAAGGCGCTGAGCATCG 420
 QY 3131 ttgggggttcagaacacgctgctggtgtgcttataatcattgaagtattatgaagaagag 3190
 DB 421 TTGGGGGTTTCAGACACGCTGCT----- 442
 QY 3191 cagcatgtaaagatgaaagacccctggcgtgagatcagatccttagaggtgaaatgag 3250
 DB 443 -----GCTGAATGCAATTG 456
 QY 3251 acttgaggaagccacacatgtagaagcaatcattgtccttgagagacagcagcagagag 3310
 DB 457 ACTTGGAAGAGCCACACATGATGAAGCAATATGTCTGTGAGACAGACGACAGAGAG 516

DB	Accession	Source	Organism	LOCUS	DEFINITION	VERSION	KEYWORDS	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
OY	3311	tcgcgcctgacactctcacaagaatgagggcccaacaaagagggaaggtgtgaaaccc	3370	1	(bases 1 to 670)	Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,Y., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.	Homo sapiens CDNA clone	Unpublished (2000)	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn	This clone is available at CHGC in Shanghai.	Location/Qualifiers	
DB	517	tcgcgcctgacactctcacaagaatgagggcccaacaaagagggaaggtgtgaaaccc	576	1	(bases 1 to 670)	Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,Y., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.	Homo sapiens CDNA clone	Unpublished (2000)	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn	This clone is available at CHGC in Shanghai.	Location/Qualifiers	
OY	3371	tcactatgagctgcagaagaacccgggaagagccctaggatlaagttatgtgt-aaa	3429	1	(bases 1 to 670)	Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,Y., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.	Homo sapiens CDNA clone	Unpublished (2000)	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn	This clone is available at CHGC in Shanghai.	Location/Qualifiers	
DB	577	tcactatgagctgcagaagaacccgggaagagccctaggatlaagttatgtgt-aaa	636	1	(bases 1 to 670)	Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,Y., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.	Homo sapiens CDNA clone	Unpublished (2000)	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn	This clone is available at CHGC in Shanghai.	Location/Qualifiers	
OY	3430	agaacgatactgagatattgttcagacacattgc	3465	1	(bases 1 to 670)	Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,Y., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.	Homo sapiens CDNA clone	Unpublished (2000)	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn	This clone is available at CHGC in Shanghai.	Location/Qualifiers	
DB	637	agaacgatactgagatattgttcagacacattgc	672	1	(bases 1 to 670)	Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,Y., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.	Homo sapiens CDNA clone	Unpublished (2000)	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn	This clone is available at CHGC in Shanghai.	Location/Qualifiers	
OY	2711	gtgagatcatataaagagcttaaacagagcataggggtgacagccagatgagactca	2770	1	(bases 1 to 670)	Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,Y., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.	Homo sapiens CDNA clone	Unpublished (2000)	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn	This clone is available at CHGC in Shanghai.	Location/Qualifiers	
DB	1	gtgagatcatataaagagcttaaacagagcataggggtgacagccagatgagactca	60	1	(bases 1 to 670)	Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,Y., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.	Homo sapiens CDNA clone	Unpublished (2000)	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn	This clone is available at CHGC in Shanghai.	Location/Qualifiers	
OY	2771	aagtcgagatcagaactagctgcagatgataaattgttgttgtaacctattgaa	2830	1	(bases 1 to 670)	Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,Y., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.	Homo sapiens CDNA clone	Unpublished (2000)	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn	This clone is available at CHGC in Shanghai.	Location/Qualifiers	
DB	61	aagtcgagatcagaactagctgcagatgataaattgttgttgtaacctattgaa	120	1	(bases 1 to 670)	Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,Y., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.	Homo sapiens CDNA clone	Unpublished (2000)	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn	This clone is available at CHGC in Shanghai.	Location/Qualifiers	
OY	2831	agttattagctcttgaaagacagaagatgacagtaaacattccattcattgtgaa	2890	1	(bases 1 to 670)	Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,Y., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.	Homo sapiens CDNA clone	Unpublished (2000)	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn	This clone is available at CHGC in Shanghai.	Location/Qualifiers	
DB	121	agttattagctcttgaaagacagaagatgacagtaaacattccattcattgtgaa	180	1	(bases 1 to 670)	Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,Y., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.	Homo sapiens CDNA clone					

QY	2951	gtctccagctcttgatgtgtcccaagctcgtggtctcccaaacggagatccatccgaaata	3010
Db	241	gctccagctcttgatgtgtcccaagctcgtggtctcccaaacggagatccatccgaaata	300
QY	3011	caagcagatcatalcaaccaccagcaatlttgctctctgatactctgaaactgccccatctcc	3070
Db	301	caagcagatcatalcaaccaccagcaatlttgctctctgatactctgaaactgccccatctcc	360
QY	3071	ctggtctggaacaacacatcagatattccaaaggcgaaaggcttggtcctgagcatcg	3130
Db	361	ctggtctggaacaacacatcagatattccaaaggcgaaaggcttggtcctgagcatcg	420
QY	3131	ttgggggttcagacacgctgctgggtgcctctattatcatgaaglttatgagaagag	3190
Db	421	ttgggggttcagacacgctgctgggtgcctctattatcatgaaglttatgagaagag	442
QY	3191	cagcatgttaaagatggaagactctgggtctgagatcagatcttgtaggtgaaatg	3250
Db	443	-----ggatgaaatg	456
QY	3251	acttagaagaagccacatgtatgaagaatcaatgctctgagacagagccacagagag	3310
Db	457	acttagaagaagccacatgtatgaagaatcaatgctctgagacagagccacagagag	516
QY	3311	tgcgcctgacactctacagagaatgagccccaataaagaagaagaatgtgtgacacc	3370
Db	517	tgcgcctgacactctacagagaatgagccccaataaagaagaagaatgtgtgacacc	576
QY	3371	tcactatgagctgcagaagaagccggagaaagccctgagatgaatgtgtgtaaa	3430
Db	577	tcactatgagctgcagaagaagccggagaaagccctgagatgaatgtgtgtaaa	635
QY	3431	gaacacatcgtgagatattgtgtcgaacattgtc	3465
Db	636	gaacacatcgtgagatattgtgtcgaacattgtc	670
RESULT 14			
LOCUS	AV646666	670 bp	mRNA
DEFINITION	AV646666	GLC Homo sapiens cDNA clone	GLCAPH08 3', mRNA sequence.
VERSION	AV646666	1	GI:9867680
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 670)		
JOURNAL	Qian,B., Wu,T., Huang,Q., Kang,B., Gao,X., Xu,Z., Xiao,H.,		
COMMENT	Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J.,		
	Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu		
	,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.		
	Homo sapiens cDNA clone		
	Unpublished (2000)		
	Contact: Zeguang Han		
	Chinese National Human Genome Center at Shanghai		
	351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai		
	201203, P. R. China		
	Tel: 86-21-50801919(ex. 45)		
	Fax: 86-21-50801922		
	Email: hanzg@chgc.sh.cn		
	This clone is available at CHGC in Shanghai.		
FEATURES	Location/Qualifiers		
SOURCE	1..670		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="GLCAPH08"		
	/clove="lib=GLC"		
	/tissue_type="corresponding non cancerous liver tissue"		
	/dev_stage="Adult"		
	/lab_host="SOLR"		

OY 3251 acttgaggaaggccacacatgatgtagaacatcatctcctgagacagagcgccacagagag 3310
||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 457 ACTTGAGAAAGGCCACACATGATGAAAGCAATCATGTCTGTGAGACGACGCCACAGAGAG 516
 Oy 3311 tgcgcctgacactctacagagatgaggcccaatacaagaaggaagtgctgacacc 3370
 Db 517 TGGCCTGACACTCTACAGAGATGAGGCCCATCCAAAGAGGAGAGTGTGTGACACCC 576
 Oy 3371 tcaactatgagctgcagaagaagccggaaagagcctagaataagtaattggtgtaaaa 3430
 Db 577 TCACTATTGAGCTGCAGAAAGCCGNNAAA-GCCTANGATTAAAGTATTGTGCT-AAA 634
 Oy 3431 gaaagatactggagtaattggtgcagacattg 3463
 Db 635 GAAAGATATCTGAGGT-TTGTGTGACACATGG 666

Search completed: July 12, 2001, 19:58:27
 Job time: 8450 sec

7